Structural and biochemical characterisation of the centromere binding factor 3

Structural and Biochemical Characterisation of the Centromere Binding Factor 3

Shomon Miah

Department of Biological Sciences
Birkbeck University

This dissertation is submitted for the degree of
Doctor of Philosophy

November 2020

I would like to dedicate this thesis to my loving parents ...

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

Shomon Miah
November 2020

Acknowledgements

Firstly I would like to thank my supervisors, Cara Vaughan and Barry Panaretou for devising this interesting project. I'd like to especially thank Cara for her mentorship, providing me with advice and support throughout

This project would not have been possible without the help of a number of people. The hard work of Wenjuan Zhang enabled the project to progress as far as it did and she provided me with a lot of guidance and showed great patience when I first started. Jonathan Lucas played a big role in this project with his work on the Ndc10 cloning, enabled us to conduct the structural and biochemical experiments. I would like to acknowledge Dr Natasha Lukoyanova, who collected all of the cryo-EM data presented in the thesis. I would like also like to thank Luis Miguel, Alex Cook, Katerina Toropova and Dave Houldershaw who advised and helped me with the cryo-EM processing. Sincere thanks to Tien-Chye Tan and Claire Bagneris and every member of the Rayne and Wolfson Laboratory who were always on hand for help with all of the experimental work.

I have been extremely fortunate to have shared my PhD experience with many wonderful and talented people. I am deeply grateful to Dina, Hugo, Alex, Joe, Emir, Trish, Josh, Fiona, Stephanie, Ugochim and especially Jenny. We have shared many fun and hilarious moments at the George and around London on weekdays, weekends and beyond. At the best and the worst of times, you guys were always there on hand to help me throughout and the PhD experience would not have been the same without you guys.

Abstract

The kinetochore is a critical multi protein complex of the mitotic machinery, which binds both the centromere DNA of the chromosome and the microtubules from the mitotic spindle. In *S. cerevisiae*, the centromere DNA is short and conserved, and associates with a specialised non-canonical nucleosome, containing a variant of histone H3 called Cse4. The centromere is specifically recognised by the kinetochore protein complex called CBF3.

The CBF3 complex is composed of a homodimer of a zinc cluster protein, Cep3, a homodimer of a DNA binding protein, Ndc10 and a heterodimer of the F-box protein Ctf13 and F-box binding protein Skp1. Genetic studies have determined that CBF3 is responsible for the recruitment of the specialised nucleosome to the point centromere, but the molecular mechanism of the recruitment is unknown.

This thesis will discuss new strategies to purify different forms of the CBF3 complex and present structural analysis conducted on a truncated CBF3 complex using single particle cryo-EM. Further work was conducted to analyse the regulation of Zn_2Cys_6 cluster domain of Cep3 DNA binding activity and whether the DNA binding domain of Ndc10 has oligomerisation activity. The results are discussed in the context of the overall implications on the current understanding of CBF3 complex function and how this adds to the understanding of kinetochore and centromere biology in budding yeast.

Table of contents

List of figures

List of tables

Nomenclature

Acronyms / Abbreviations

Å Angstrom

λ-PP Lambda protein phosphatase

μg Microgram

μL Microlitre

μM Micromolar

APC/C Anaphase-Promoting Complex/Cyclosome

bp Base pair

C. elegans *Caenorhabditis elegans*

CBF3 Centromere binding factor 3

CBP Calmodulin binding peptide

CDEI Centromere determining element I

CDEII Centromere determining element II

CDEIII Centromere determining element III

CEN3 Centromere of chromosome 3

CEN Centromere

cryo-EM Cryogenic electron microscopy

CTF Contrast Transfer Function

Da Dalton

DNA Deoxyribonucleic acid

dNTPs Deoxynucleotide triphosphates

DTT Dithiothreitol

E. coli *Escherichia coli*

EDTA DescriptionEthylenediaminetetraacetic acid

EMSA Electrophoretic mobility shift assay

g Gram

HIS Histidine

IMAC Immobilized metal affinity chromatography

IPTG Isopropyl-β-D-1-thiogalactopyranoside

K. lactis *Kluyveromyces lactis*

Kb Kilobase

kDa Kilodalton

LB Lysogeny broth

L Litre

LRR Leucine rich repeats

mA Milliamps

Mb Megabase

MCC Mitotic Checkpoint Complex

Min Minute

mL Millilitre

mM Millimolar

NEB New England Biolabs

ng Nanogram

OD Optical density

PCR Polymerase chain reaction

PDB Protein Data Bank

S. cerevisiae *Saccharomyces cerevisiae*

SAC Spindle Assembly Checkpoint

SDS-PAGE Sodium dodecyl sulfate–polyacrylamide gel electrophoresis

Sec Second

SGD Stochastic gradient descent

TAE Tris-Acetic acid - EDTA

TBE Tris-Borate-EDTA

TB Terrific Broth

TEV Tobacco Etch Virus

V Voltage

YEPD Yeast extract peptone dextrose

Chapter 1

Introduction

1.1 Chromosome Segregation

Spatial and temporal regulation of chromosome segregation during cell division is critical to maintaining cellular homeostasis. In mammalian cells, at the mitosis stage of the cell cycle (Fig 1.1), chromosomes are condensed, the nuclear envelope breaks down and the sister chromatids establish attachments to the microtubules from the mitotic spindle (Mitchison and Salmon, 2001). The chromosomes align at the metaphase plate, before separating to the mother and daughter cells (Mitchison and Salmon, 2001). After separation, the nuclear envelope reforms around the chromosomes (Mitchison and Salmon, 2001). There are a number distinct differences in the process of mitosis between different eukaryotic organisms and therefore this thesis will primarily discuss the molecular mechanisms of mitosis in mammalian and budding yeast cells.

To accurately segregate the chromosomes in budding, the cell utilises many different proteins and protein complexes, to enable and regulate the process, establishing a level of complexity, which is reflected in the need for fidelity (Cheeseman and Desai, 2008, Cheeseman, 2014). The breakdown of these regulatory mechanisms can be attributed to many different disease states, including tumour progression in leukaemia, breast and lung cancer as well as Down's syndrome and Alzheimer's disease (Bharadwaj and Yu, 2004, Antonarakis et al., 2004, Weaver and Cleveland, 2006, Holland and Cleveland, 2012, Gordon et al., 2012, Yurov et al., 2014). Many of the disease states are associated with cells becoming aneuploid, where the cells contain an abnormal

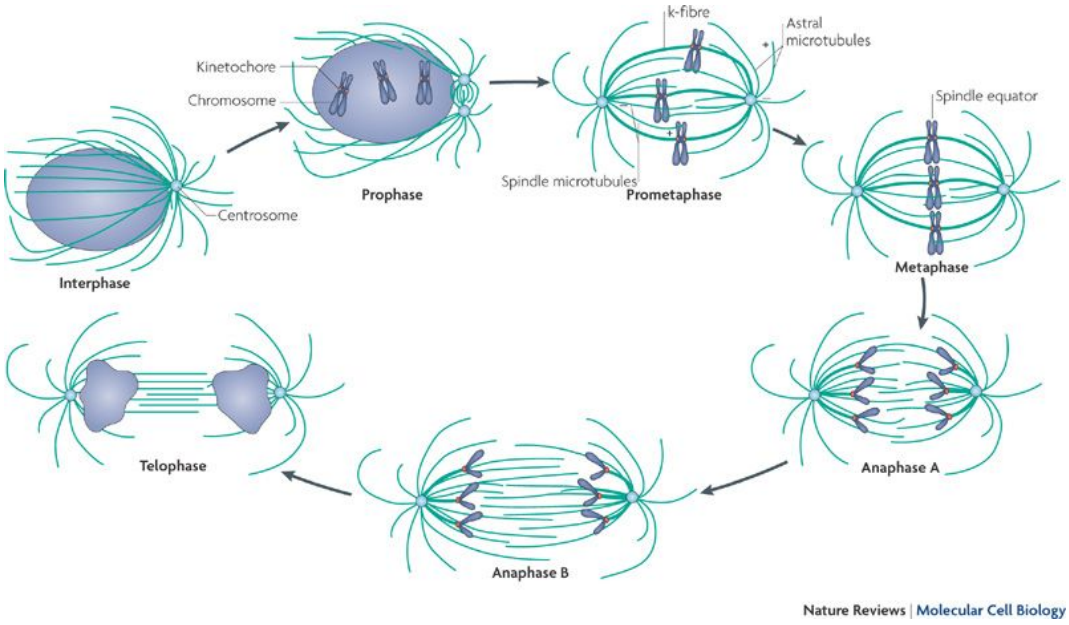


Fig. 1.1 **Stages of mitosis.** Illustration of the different stages of mitosis in mammalian cells and the organisation of the chromosomes and the microtubules. The microtubules are highlighted in green and the chromatids are coloured in blue. This figure was taken from (Walczak et al., 2010).

number of chromosomes (Holland and Cleveland, 2012, Gordon et al., 2012). Aberrant chromosome segregation can arise through a number of different ways, which includes break down in regulation at the spindle assembly checkpoint (SAC) and improper attachment of chromatids to the microtubules via the large protein complex called the kinetochore (Gordon et al., 2012, Bharadwaj and Yu, 2004, Kops et al., 2005, Thompson et al., 2010).

To achieve accurate chromosome segregation in budding yeast, the kinetochore must recognise a specific sequence of DNA at a specific locus called the centromere (CEN) (Cooke et al., 1993), the kinetochore attachment to the mitotic spindle must be robust to establish a physical link between the centromere DNA and the microtubules, and the attachment to the mitotic spindle must satisfy the requirements of the SAC (McIntosh et al., 2002). The SAC functions to prevent the onset of anaphase when the chromosomes are unattached or improperly attached to the mitotic spindle (Vleugel et al., 2012). This results in the cells avoiding alteration in chromosome copy numbers, and the subsequent results of that (Santaguida and Amon, 2015). In this project the

centromere is defined as the DNA sequence on which the kinetochore assembles and the kinetochore refers to the protein assembly which is able to assemble on the centromere.

1.2 The Kinetochore

The kinetochore is a critical component of the mitotic machinery whose assembly is regulated during the different stages of the cell cycle and the timing of the assembly is species dependent (Cheeseman and Desai, 2008, Cheeseman, 2014). In human and yeast, during prometaphase, the kinetochore forms interactions with microtubules to align the chromosomes along the metaphase plate (Magidson et al., 2011). At metaphase, the kinetochore facilitates chromatid interaction with microtubules from opposite spindle poles. At anaphase, the kinetochore locates to opposite spindle poles, to drive chromatid separation at anaphase (Fig 1.2). The outer kinetochore also plays a crucial role in activating the SAC function in both mammalian and budding yeast cells (Aravamudhan et al., 2015, Joglekar, 2016).

1.2.1 Architecture and function of the kinetochore

The kinetochore is composed of a large number of different proteins, the majority of which are conserved between different eukaryotic organisms (Cheeseman and Desai, 2008, Joglekar et al., 2008, 2009, Santaguida and Musacchio, 2009, Cheeseman, 2014). Structural and biochemical studies into the higher eukaryotic kinetochores have proven to be challenging. This is attributed to the size of the centromere DNA, which can be up to several megabases in length, and the increased DNA complexity leads to increased protein complexity and higher eukaryotes have a multi-kinetochore model with increased number of copies of individual proteins, where each kinetochore forms multiple interactions with microtubules (Harrington et al., 1997, Yang et al., 2000, McEwen et al., 2001). As a result, the budding yeast, *Saccharomyces cerevisiae* (*S. cerevisiae*) has become the model organism of choice for studies involving reconstitution of the kinetochore complex and structural studies because it is likely the most structurally compact kinetochore with the shortest centromere of the eukaryotes forming a simple single kinetochore unit and as a result can be used to provide information on kinetochore assembly in higher eukaryotes (Lang et al., 2018, Leber et al., 2018, Zhang et al., 2018). The following description focuses on the *S. cerevisiae* kinetochore protein complex

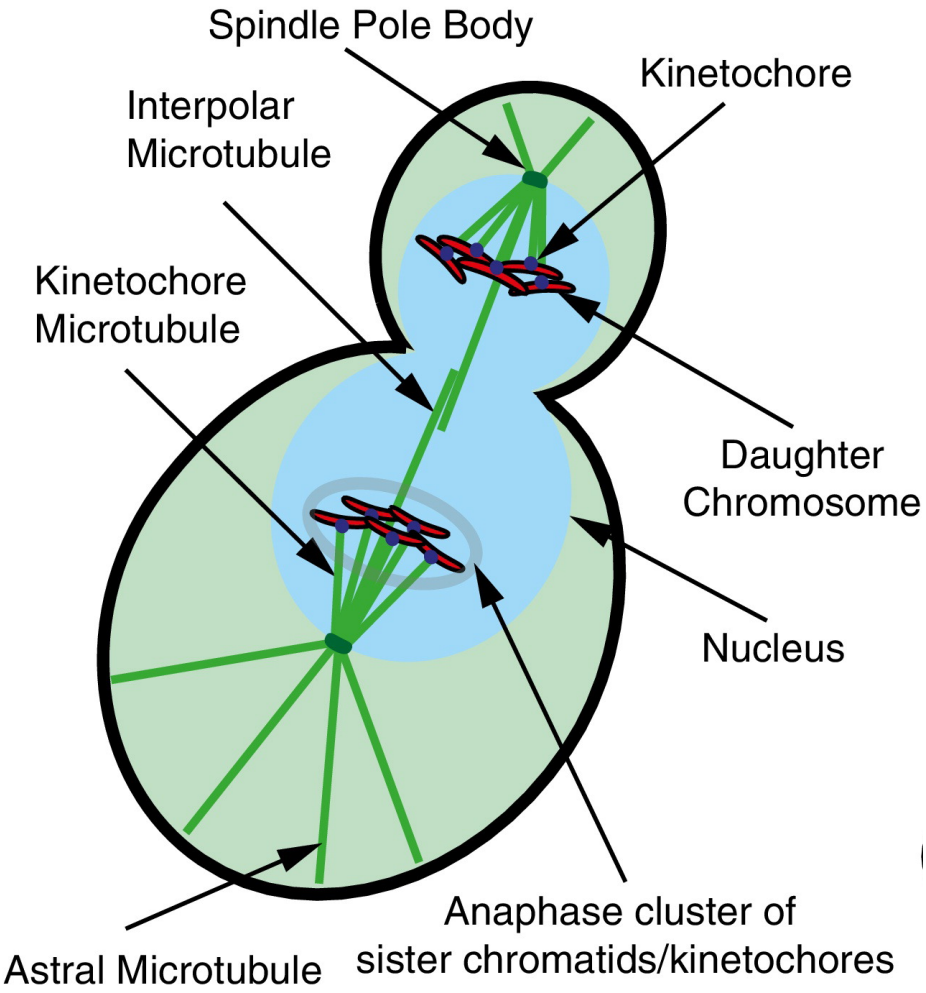


Fig. 1.2 **Kinetochore attachments to chromosomes.** This illustration describes the interaction between microtubules, the kinetochore and the chromosomes at anaphase in budding yeast. The kinetochore is highlighted in blue, the chromosome in red and microtubules in green. This figure was taken from (Cieśliński and Ries, 2014)

unless otherwise noted. The protein complex names are capitalised to distinguish them from the names of the individual proteins.

The core of the *S. cerevisiae* kinetochore can be differentiated into the outer kinetochore, and the inner kinetochore (Fig 1.3) (Cheeseman, 2014, Cieśliński and Ries, 2014). The outer kinetochore is responsible for the affinity and interaction of the kinetochore with microtubules (Wigge and Kilmartin, 2001, Wei et al., 2005, Cheeseman et al., 2006,

Joglekar et al., 2006, Lampert et al., 2010, Alushin et al., 2010). The inner kinetochore facilitates the deposition of centromere-specific nucleosomes to the centromere and provides a platform to assemble the rest of the kinetochore (Fig 1.3) (Cooke et al., 1993, Sorger et al., 1994, He et al., 2001).

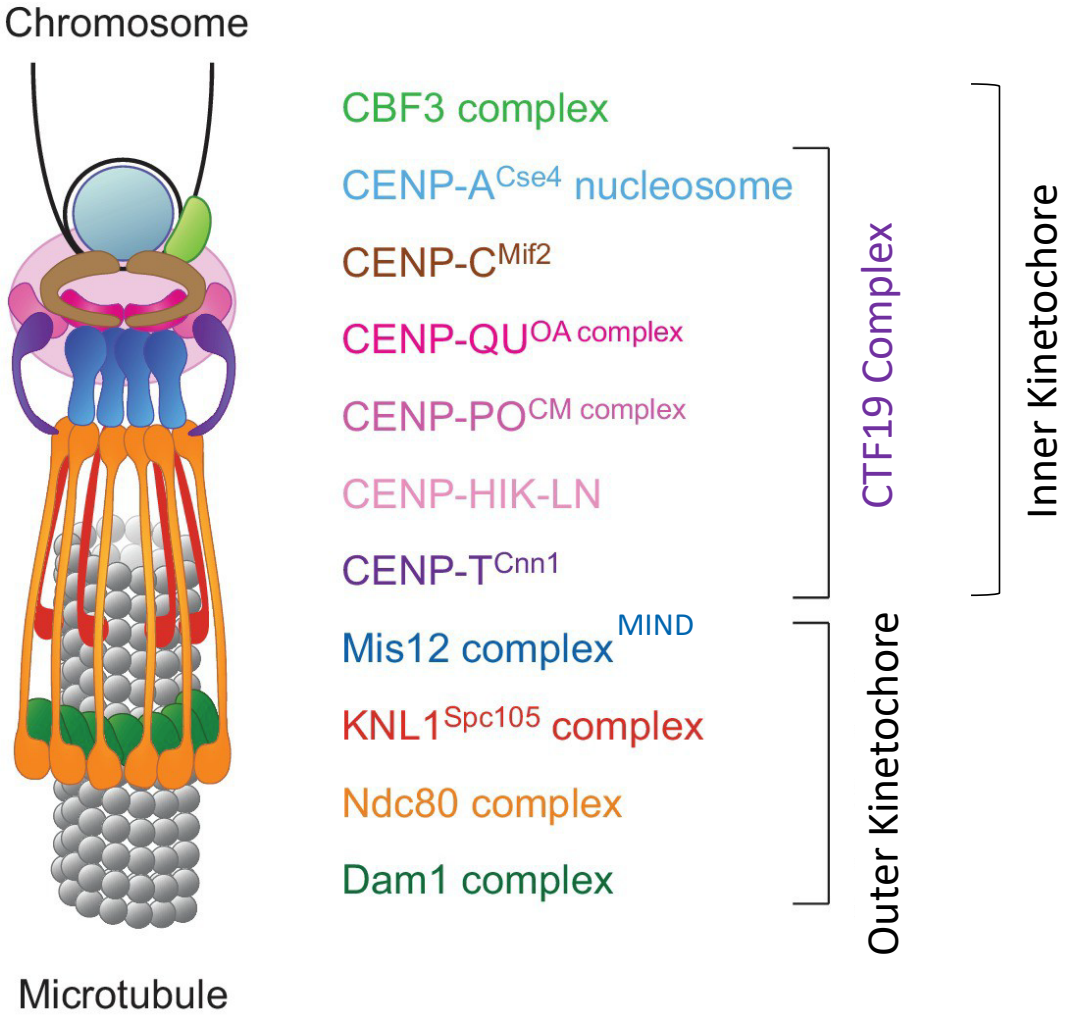


Fig. 1.3 **Composition of the *S. cerevisiae* Kinetochore.** The diagram describes the organisation of the kinetochore complexes relative to one another. The diagram illustrates the outer kinetochore and the inner kinetochore complexes of CTF19 and CBF3. This figure was taken and adapted from (Lang et al., 2018).

S. *cerevisiae* Protein complex	Human homolog	S. *cerevisiae* protein subunit	Human homolog
DAM1	-	Ask1	-
		Dad1	-
		Dad2	-
		Dad3	-
		Dad4	-
		Dam1	-
		Duo1	-
		Spc34	-
		Spc19	-
		Hsk3	-
SPC105	KNL1	Spc105	Knl1
		Kre28	Zwint
NDC80	HEC1	Ndc80	Hec1
		Nuf2	Nuf2
		Spc24	Spc24
		Spc25	Spc25
MIND	MIS12	Mtw1	Mis12
		Dsn1	Dsn1
		Nnf1	Pmf1
		Nsl1	Nsl1
CTF19	CCAN	Ctf19	CENP-P
		Okp1	CENP-Q
		Mcm21	CENP-0
		Ame1	CENP-U
		Chl4	CENP-N
		Iml3	CENP-L
		Mcm22	CENP-K
		Mcm16	CENP-H
		Nkp1	-
		Ctf3	CENP-I
		Nkp2	-
		Mif2	CENP-C
		Cnn1	CENP-T
		Wip1	CENP-W
Cse4 Nucleosome	CENP-A Nucleosome	Cse4	CENP-A
		H2A	H2A
		H2B	H2B
		H4	H4
CBF3	-	Cep3	-
		Ctf13	-
		Skp1	-
		Ndc10	-

Table 1.1 **List of Kinetochore proteins.** Details on the protein complexes and the individual subunits of the budding yeast kinetochore complexes and their homologs in human cells.

1.2.2 The outer kinetochore

The outer kinetochore proteins provide specificity and affinity to facilitate the interaction between the kinetochore and the microtubules in most eukaryotic organisms (Wigge and Kilmartin, 2001, Wei et al., 2005, Cheeseman et al., 2006, Joglekar et al., 2006, Alushin et al., 2010, Lampert et al., 2010). These proteins assemble on a platform provided by the inner kinetochore proteins when the cell transitions from G2 to mitosis in human cells (Gascoigne and Cheeseman, 2013, Cheeseman, 2014). These protein complexes then disassemble when the cell exits from mitosis at anaphase and telophase (Gascoigne and Cheeseman, 2013, Cheeseman, 2014). The *S. cerevisiae* outer kinetochore is composed of a number of regulatory proteins as well as three core protein assemblies, the heterotetrameric NDC80 protein complex, the heterodecameric DAM1 complex and the SPC105 complex (Fig 1.3) (De Wulf et al., 2003, Biggins, 2013). These core protein assemblies maintain stable association of the kinetochore with the microtubule plus end, as it lengthens and shortens and maintains association with the inner kinetochore (Cheeseman et al., 2006).

The NDC80 protein complex is an end-to-end rod-like structure composed of two heterodimers of Ndc80 – Nuf2 and Spc24 – Spc25 complexes (Fig 1.4.A) (Wei et al., 2005, Ciferri et al., 2005). The NDC80 complex is required for the recruitment of the DAM1 protein complex to the microtubule (Janke et al., 2002, Tien et al., 2010, Lampert et al., 2010, 2013). The NDC80 complex also acts to form two kinetochore receptors for additional kinetochore attachment (Malvezzi et al., 2013). NDC80 is able to form a receptor by associating with SPC105 complex and the MIND complex to form the KMN network (Wei et al., 2006, Petrovic et al., 2010). NDC80 forms a second receptor by associating with proteins CENP – TWSX, at the same interaction site as the MIND complex (Schleiffer et al., 2012, Bock et al., 2012, Nishino et al., 2013, Rago et al., 2015).

The heterodecameric DAM1 protein complex forms a ring structure, composed of 17 copies of the decamer (Fig 1.4.B) (Jenni and Harrison, 2018). The ring encircles the microtubules and it has been postulated that the complex is able to slide down but remain attached to the microtubule during depolymerization (Miranda et al., 2005, Asbury et al., 2006, Hanisch et al., 2006, Westermann et al., 2006, Grishchuk et al., 2008). In addition to DAM1 and NDC80 protein complexes interacting with the microtubules, they also make contacts with each other and these contacts are regulated by phosphorylation, which suggests that the assembly of the two complexes is

temporally regulated to enable it to couple the microtubules to chromosome segregation (Cheeseman et al., 2002, Gestaut et al., 2008, Kim et al., 2017, Jenni and Harrison, 2018).

The SPC105 protein complex is a highly elongated assembly, which binds to microtubules with low affinity (Pagliuca et al., 2009). The recruitment of the complex is only dependent on the CBF3 complex (Pagliuca et al., 2009). SPC105 is not required for binding to microtubules but instead it recruits microtubule-associated proteins and kinesins to the kinetochore, to ensure bi-orientation of sister chromatids on the mitotic spindle (Pagliuca et al., 2009). The SPC105 complex is also essential for recruitment of SAC proteins to the kinetochore (Pagliuca et al., 2009).

The outer kinetochore also includes a number of regulatory proteins which interact transiently such as Ipl1 kinase, Pp1 phosphatase and regulatory kinesins such as Kip1, Cin8 and Kip3 to facilitate and regulate kinetochore function (Tytell and Sorger, 2006, Liu et al., 2010, Dimitrova et al., 2016). The outer kinetochore also plays a crucial role in activating the SAC function. Studies into budding yeast and metazoan SAC have uncovered the mechanistic detail on the role of the kinetochore in SAC function (Aravamudhan et al., 2015, Joglekar, 2016). The architecture of the outer kinetochore acts as a mechanical switching, turning on or silencing the SAC. This mechanical switching is regulated by the Mps1 kinase (Aravamudhan et al., 2015). Mps1 is bound by Ndc80 protein of the outer kinetochore, at a binding site overlapping with the microtubule binding site (Ji et al., 2015). When the kinetochore is unattached or improperly attached, Mps1 binds to Ndc80, and phosphorylates its target site on Spc105, which initiates a cascade of biochemical activity, which results in the recruitment of SAC proteins to the kinetochore, including the components of the Mitotic Checkpoint Complex (MCC)(Joglekar, 2016). The MCC inhibits the Anaphase Promoting Complex/Cyclosome (APC/C), an E3 ubiquitin ligase that targets Securin and Cyclin B for degradation, preventing metaphase-anaphase transition, until the kinetochore attachment is corrected (Izawa and Pines, 2015).

1.2.3 The MIND complex

The MIND complex is a heterotetrameric protein complex composed of the proteins Mtw1, Nnf1, Nsl1 and Dsn1 (Fig 1.5) (Maskell et al., 2010, Hornung et al., 2011, Dimitrova et al., 2016). The complex acts as a hub, which bridges the centromere-

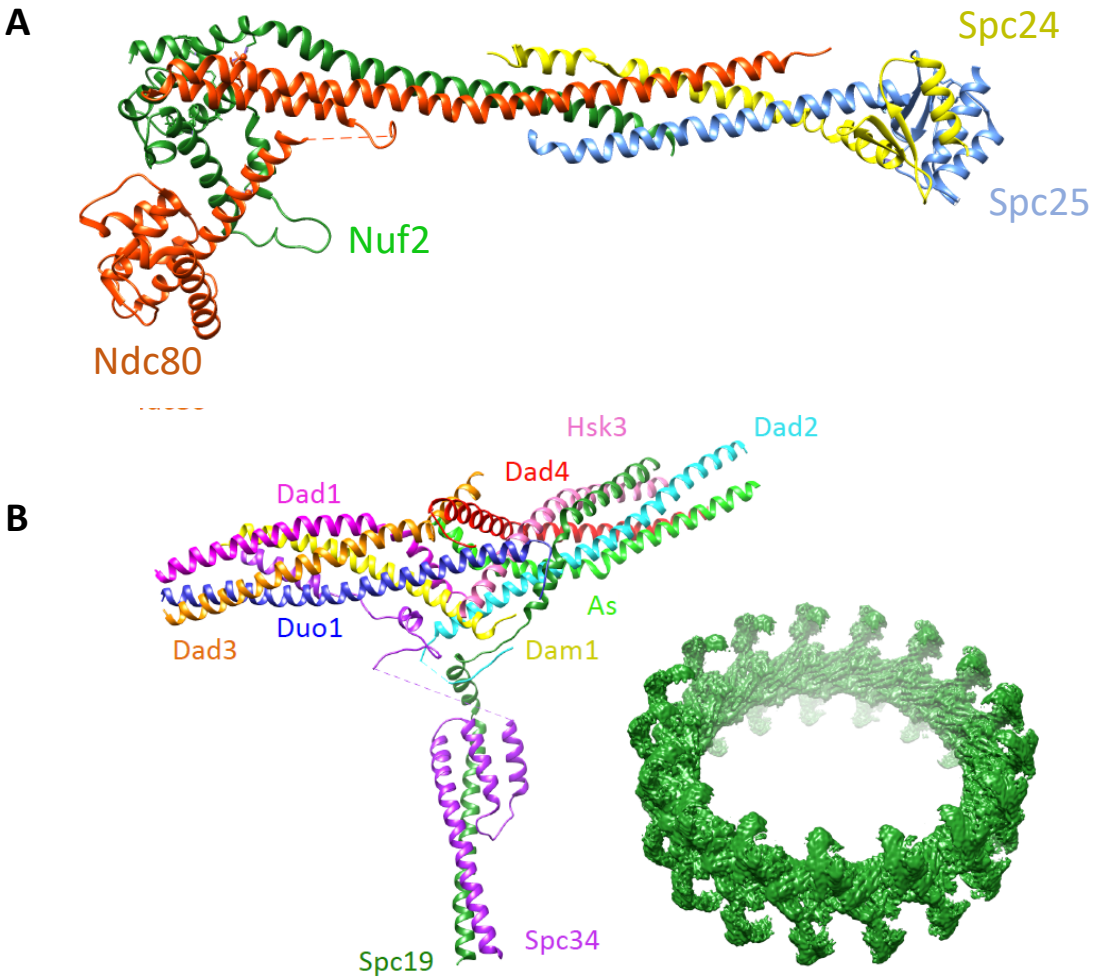


Fig. 1.4 **Structures of the outer kinetochore components.** (A) 2.8 Å crystal structure of NDC80 tetramer, with the individual subunits labelled. PDB: 5TD8. (B) 4.5 Å cryo-EM model of a single DAM1 complex (left) and a 6.5 Å cryo-EM map of the DAM1 complex ring. PDB: 6CFZ. EMDB: 7446.

associated part of the kinetochore to the outer microtubule associated region (Fig 1.3) (De Wulf et al., 2003, Biggins, 2013). The complex is essential for bi-orientation of sister chromatids and the subsequent accurate segregation during mitosis (Cheeseman et al., 2006, Kline et al., 2006, Pagliuca et al., 2009, Hornung et al., 2014).

The MIND complex forms a 200 Å Y-shaped rod, with two globular heads projecting from a long coiled coil shaft (Fig 1.5) (Corbett et al., 2010, Dimitrova et al., 2016). At

one end of the rod, the MIND complex interacts with NDC80 and SPC105 (Dimitrova et al., 2016). At the other end, the complex interacts with Mif2 and the COMA complex (Hornung et al., 2014, Dimitrova et al., 2016). The function of the MIND complex is regulated by phosphorylation (Emanuele et al., 2008, Yang et al., 2008, Akiyoshi et al., 2013, Kim and Yu, 2015). In budding yeast, this is conducted by the kinase Ipl1 (Aurora B in mammalian cells) (Biggins and Murray, 2001, Stern and Murray, 2001, Cheeseman et al., 2002, Tien et al., 2010, Krenn and Musacchio, 2015). Ipl1 promotes the recruitment of the MIND complex to the centromere by phosphorylating the Dsn1 subunit at two conserved sites (Akiyoshi et al., 2013, Dimitrova et al., 2016). The phosphorylation relieves auto-inhibition of Dsn1, enabling the MIND complex to interact with Mif2 and the COMA complex (Dimitrova et al., 2016).

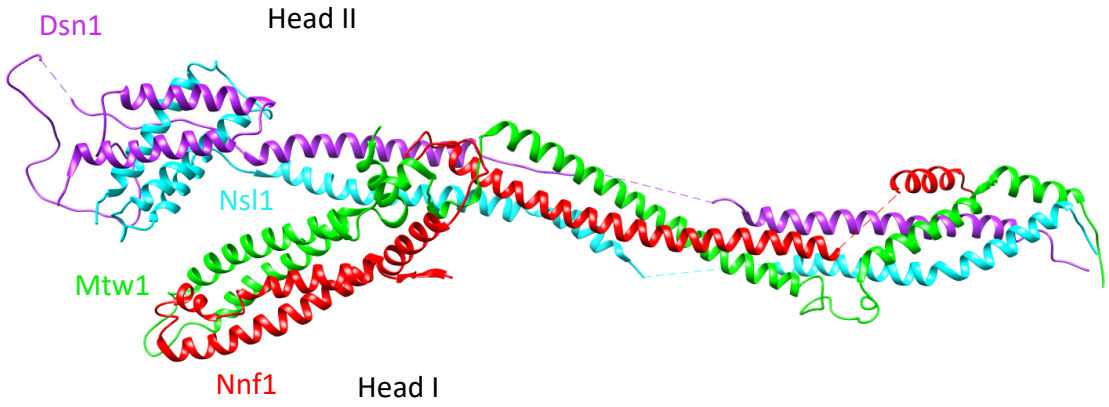


Fig. 1.5 **Crystal structure of the MIND complex.** 3.2 Å crystal structure of MIND complex from *K. lactis*. The subunits are labelled and colour coded with the appropriate polypeptide chain. PDB: 5T58.

1.2.4 CTF19 complex

The CTF19 complex, homologous to the constitutive centromere – associated network (CCAN) in higher eukaryotes, is a 13-subunit protein complex, which is able to associate with Mif2 and the Cse4 nucleosome (Fig 1.6) (Ortiz et al., 1999, Pekgöz Altunkaya et al., 2016). The CTF19 complex is responsible for localising and forming an assembly scaffold for the outer kinetochore by associating with the Cse4 nucleosome (Pekgöz Altunkaya et al., 2016, Fischböck-Halwachs et al., 2019). Two copies of the CTF19

protein complexes bind to a nucleosome (Weir et al., 2016). The 13 subunits of the CTF19 complex forms a defined structure which can be divided into sub-complexes: Mif2, COMA, CTF3, CHl4 and CNN1 (Biggins, 2013).

The protein Mif2 is an essential protein which associates with the CTF19 complex and shares homology with the mammalian centromere protein CENP-C (Brown, 1995, Meluh and Koshland, 1995). Two copies of Mif2 binds Cse4 nucleosomes with high affinity and selectivity (Xiao et al., 2017). Mif2 has been localised to a 30 bp stretch of CDEII (Xiao et al., 2017). Mif2 also forms interactions with Ame1 and Okp1 of the COMA complex to form a binding scaffold to bind Mtw1 of the MIND complex (Hornung et al., 2014). The formation of the binding scaffold and the association of the MIND complex enable the formation of the outer kinetochore (Hornung et al., 2014)

The COMA complex, composed of the proteins: Ctf19, Okp1, Mcm21 and Ame1 (Fig 1.6), is responsible for linking the inner kinetochore to the outer kinetochore (De Wulf et al., 2003, Musacchio and Desai, 2017). The Okp1 and Ame1 subunits are essential to kinetochore function, as they bridge the CTF19 complex to the outer kinetochore (Ortiz et al., 1999, Hornung et al., 2014, Dimitrova et al., 2016, Musacchio and Desai, 2017). The Okp1 and Ame1 subunits also facilitates downstream recruitment of CTF19 components (Pekgöz Altunkaya et al., 2016). The Ctf19-Mcm21 subunits form a ridged Y-shaped scaffold which associates with Okp1-Ame1 and interacts with Ctf3, to localise the Ctf3 sub-complex to CTF19 (Pot et al., 2002, Schmitzberger and Harrison, 2012, Hinshaw and Harrison, 2019).

The protein Chl4 forms a heterodimer with the protein Iml3 to form a platform to recruit further CTF19 components, including subunits of the COMA complex and CTF3 complex (Guo et al., 2013, Hinshaw and Harrison, 2013, 2019). The exposed surface of an extended β-sheet of Iml3 provides positively charged residues, positioned to potentially interact with DNA (Hinshaw and Harrison, 2019).

The CTF3 complex is composed of the proteins, Ctf3, Mcm16 and Mcm22 (Pekgöz Altunkaya et al., 2016). CTF3 associates with the CTF19 complex by forming contacts with Mcm21 of the COMA complex and Iml3 (Hinshaw and Harrison, 2019). The CTF3 complex also associates with the heterodimer of Cnn1 and Wip1, and is responsible for the localisation of CNN1 complex to the yeast centromere (Pekgöz Altunkaya et al., 2016). The localisation is facilitated by the interactions between Cnn1 and the heterodimer of Mcm16-Mcm22 (Pekgöz Altunkaya et al., 2016). These interactions

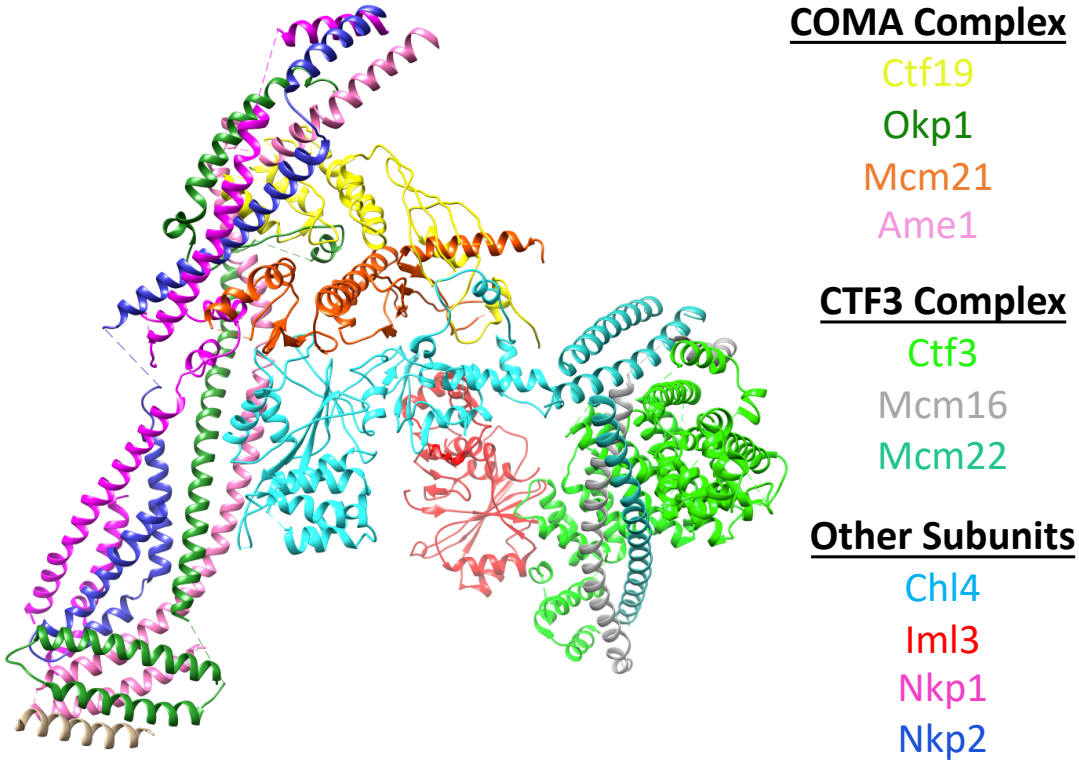


Fig. 1.6 **Structure of CTF19 complex.** 4.25 Å cryo-EM structure of the CTF19 complex, without Mif2 and the sub-assembly of Cnn1 and Wip1. The individual subunits are coloured, and grouped by their sub-complexes. The sub-complexes are highlighted in bold and underlined. PDB: 6NUW.

position the assembly, specifically the Cnn1 N-terminal extension, such that the interaction site can recruit two NDC80 molecules (Pekgöz Altunkaya et al., 2016).

1.3 The Centromere

The centromere DNA facilitates the separation of chromosomes during mitosis and is required for association with the mitotic spindle in budding yeast (Vig, 1982, Bernard et al., 2001, Nasmyth, 2002). The DNA sequence provides a platform for the inner kinetochore proteins, to establish and assemble the kinetochore (Fig 1.3) (Lawrimore et al., 2011, Fukagawa and Earnshaw, 2014, Yamagishi et al., 2014, Thakur et al., 2015, Musacchio and Desai, 2017).

Most eukaryotes have monocentric chromosomes where the centromere is localised to a single region (McKinley and Cheeseman, 2016). There are two types of localised centromeres called regional centromeres and point centromeres (Wevrick and Willard, 1989, Clarke, 1990, Hegemann and Fleig, 1993, Pluta et al., 1995, Ngan and Clarke, 1997). Regional centromeres can vary in size from several Kb to several Mb in length, composed of a hierarchical array of α-satellite repeats, or non-repetitive sequences flanked by less ordered inverted repeats and the DNA sequence in both examples, contributes but is not sufficient for centromere function (Wevrick and Willard, 1989, Pluta et al., 1995, McKinley and Cheeseman, 2016). These are commonly found in mammals and fission yeast organisms (Fishel et al., 1988, Mitchell, 1996). Point centromeres are typically found in budding yeast organisms such as *Saccharomyces cerevisiae*, where the centromere is approximately $\approx 125 - 150$ bp in length (Fig 1.7) and the DNA sequence is sufficient for centromere and kinetochore function (Lechner and Carbon, 1991, Pluta et al., 1995, Clarke, 1998). Some plants and the nematode *Caenorhabditis elegans* have holocentric chromosomes, where the centromere sequences span the length of the chromosome and interact with spindle microtubules at each of the sites (Albertson and Thomson, 1982, Maddox et al., 2004, Melters et al., 2012, Steiner and Henikoff, 2014).

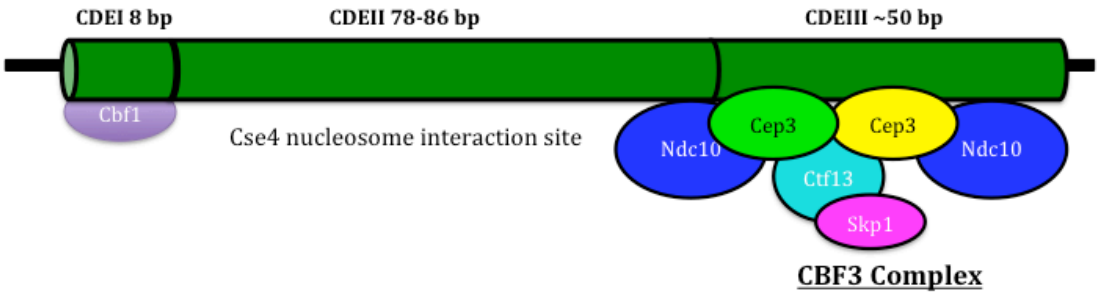


Fig. 1.7 **Point Centromere.**Illustration of the point centromere and the three DNA elements it is composed of. The diagram also shows the leucine zipper protein Cbf1 bound to CDEI and the CBF3 complex bound to CDEIII.

The *S. cerevisiae* centromere was the first to be isolated and studied (Clarke and Carbon, 1980, Bloom and Carbon, 1982, Clarke, 1998). The point centromere sequence is highly conserved, common to all chromosomes and composed of three distinct regions called centromere-determining elements (CDE) (Clarke and Carbon, 1980, Bloom and Carbon, 1982, Clarke, 1998). The first element, Centromere Determining Element I (CDEI) is a conserved 8 base pair palindrome which serves as the binding site for the

non-essential dimeric leucine rich zipper protein, Cbf1 (Cai and Davis, 1989, Baker et al., 1989, Cai and Davis, 1990). When the CDEI sequence is deleted, the rate of chromosome loss is increased 10 – fold, suggesting that the sequence is not essential but it does have an important role in centromere function (Ng and Carbon, 1987). The second element, CDEII, is 78-86 base pairs in length where 90% of its composition is Adenine and Thymine. This region of the centromere is essential, and mutations to the sequence have major effects on centromere function (Gaudet and Fitzgerald-Hayes, 1987, Spencer and Hieter, 1992, Henikoff and Furuyama, 2012, Krassovsky et al., 2012). The sequence is not conserved but the base composition is (Baker and Rogers, 2005). At present, no direct binding partners for CDEII have been found. It has been speculated that the CDEII may interact with the A-T hook motif of Mif2, it could bind to Ndc10 without the rest of CBF3 subunits or the CDEII region may loop out to interact with the Cse4 nucleosome (Brown et al., 1993, Meluh and Koshland, 1995, Pietrasanta et al., 1999, Espelin et al., 2003, Díaz-Ingelmo et al., 2015). The third element (CDEIII) is approximately 25 – 50 base pairs in length and serves as the binding site for the essential Centromere Binding Factor 3 (CBF3) protein complex (Jehn et al., 1991, Lechner and Carbon, 1991, Clarke, 1998, Krassovsky et al., 2012). The CDEIII element displays pseudo dyad symmetry composed of two triplet-half sites, CCG and TGT, responsible for binding CBF3 (Fitzgerald-Hayes et al., 1982). Mutations to CDEI and CDEII sequence decreases fidelity of chromosome segregation and when CDEIII sequence is mutated, the centromere function is lost and the kinetochore can't assemble, indicating its essential nature (Carbon and Clarke, 1984, Ng and Carbon, 1987).

1.4 The Cse4 nucleosome

In most eukaryotes, centromere identity and propagation is determined by specialised nucleosomes (McKinley and Cheeseman, 2016). These nucleosomes wrap less DNA than the conventional nucleosomes and contain a variant histone H3 called CENP-A in humans (Heslop-Harrison and Schwarzacher, 2013, Camahort et al., 2009, Schalch and Steiner, 2017). Budding yeast have a homolog of this protein called Cse4, and genetic studies have shown Cse4 can rescue centromere function when CENP-A function in mammalian cells is removed (Meluh and Koshland, 1997, Wieland et al., 2004). The specialised nucleosomes are deposited at the centromere during the G1 phase of the cell cycle and its deposition is dependent on the histone chaperone HJURP (Jansen et al., 2007, Dunleavy et al., 2009, Foltz et al., 2009, Barnhart-Dailey et al., 2017).

HJURP also has a homolog in budding yeast called Scm3, responsible for localising Cse4 nucleosomes to the centromere (Mizuguchi et al., 2007, Stoler et al., 2007, Camahort et al., 2007, Sanchez-Pulido et al., 2009, Pidoux et al., 2009, Williams et al., 2009). Scm3 is specific to Cse4 and cannot recognise histone H3 (Cho and Harrison, 2011). Localisation of Scm3 and in turn the Cse4 nucleosome is dependent on the CBF3 Complex, (Section 1.5) (Camahort et al., 2007). The conservation of the Scm3 protein between mammalian cells and budding yeast suggests the mechanism of deposition of the centromere specific nucleosome could be conserved between the eukaryotes. In budding yeast Scm3 localisation and hence the localisation and deposition of Cse4 nucleosome is dependent on Ndc10 subunit of the CBF3 complex (Camahort et al., 2007, Cho and Harrison, 2011, 2012).

1.5 The Centromere Binding Factor 3 Protein Complex

The Centromere Binding Factor 3 Complex (CBF3 Complex) is an essential *S. cerevisiae* protein complex, unique to budding yeast, composed of four subunits: homodimer of Cep3, heterodimer of Ctf13 - Skp1 and a homodimer of Ndc10 (Lechner and Carbon, 1991). The protein complex is essential to budding yeast viability and when removed from the cell, the observed phenotypes include loss of outer kinetochore – microtubule attachment, loss of kinetochore assembly and inhibition of centromere specific nucleosome deposition (Sorger et al., 1994, Camahort et al., 2007, Lang et al., 2018). Two forms of the complex have been identified, a three subunit complex composed of Cep3, Skp1 and Ctf13, called the core CBF3 complex and the full CBF3 complex composed of the core complex and Ndc10 (Fig 1.8) (Lechner and Carbon, 1991, Doheny et al., 1993, Lechner, 1994, Strunnikov et al., 1995, Stemmann and Lechner, 1996, Espelin et al., 1997, Russell et al., 1999, Leber et al., 2018, Zhang et al., 2018). The Protein - DNA crosslinking experiment also suggested an extended CBF3 complex can form when two copies of Ndc10 associate with the full CBF3 complex (Espelin et al., 1997).

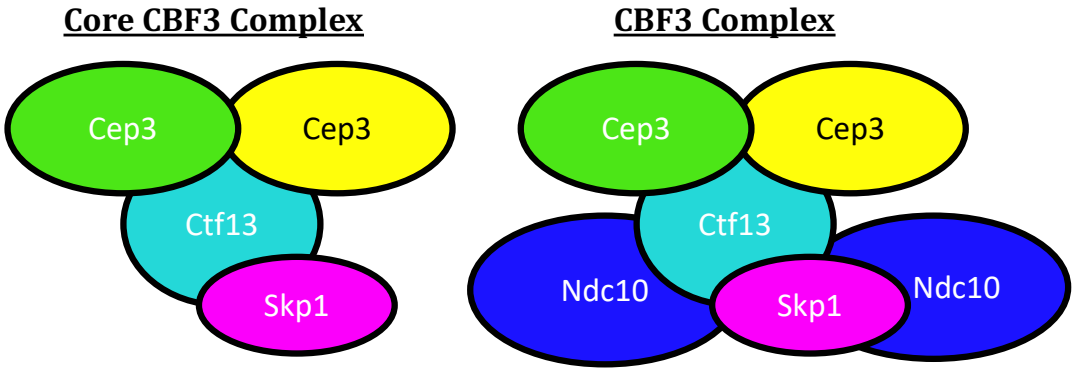


Fig. 1.8 **Two forms of the CBF3 complex.** Illustration of the two forms of the CBF3 complex identified. The core complex contains three subunits and the complex is completed by the addition of Ndc10 to form the CBF3 complex.

1.5.1 Functions of the CBF3 complex

The CBF3 complex has multiple functions in the cell. The primary function of the CBF3 complex is to facilitate the deposition of Cse4 containing nucleosomes to the point centromere for epigenetic specification of the DNA sequence (Furuyama and Biggins, 2007, Cole et al., 2011, Black et al., 2010, Krassovsky et al., 2012). The recruitment is facilitated in tandem by the Ndc10 subunit and the histone chaperone Scm3 (Camahort et al., 2007, Shivaraju et al., 2011). Scm3 is selective for Cse4 nucleosomes because it can form specific interactions with the histone fold domain of Cse4 and histone H4 (Black et al., 2004, Cho and Harrison, 2011).In *Kluyveromyces. lactis*(*K. lactis*), the Scm3 residues 41-115 at the N-terminus are responsible for this selective activity (Fig 1.9) (Cho and Harrison, 2011). Scm3 also forms a direct interaction with the CBF3 complex, by associating with the Ndc10 subunit (Cho and Harrison, 2012). In *K. lactis* the Scm3 interaction site is located at the C-terminal domains 4 and 5 of Ndc10, which are able to interact with N-terminal residues 1-28 of Scm3 (Fig 1.9) (Cho and Harrison, 2012). The Ndc10 and Cse4 interaction sites as a result don't overlap and can occur simultaneously (Cho and Harrison, 2012).

In addition to recruiting the Cse4 nucleosome to the centromere DNA, the CBF3 complex is also required to initiate kinetochore assembly (Sorger et al., 1994, Poddar et al., 2004, Lang et al., 2018). This has been investigated *in vivo* using *ndc10* mutants (He et al., 2001, Poddar et al., 2004, Lang et al., 2018). It has now been established

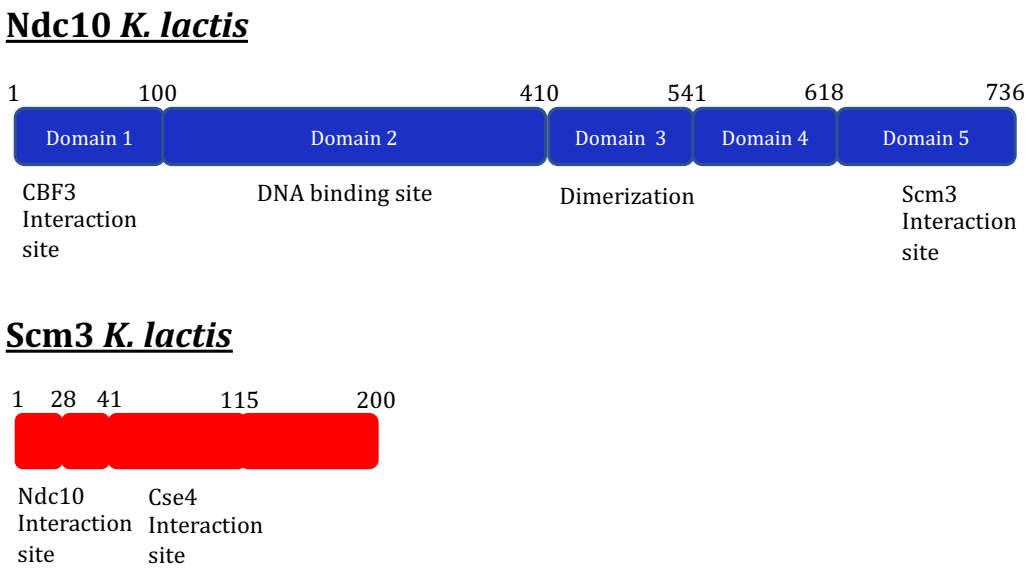


Fig. 1.9 **Ndc10 and Scm3.** Illustration of the domain organisation of Ndc10 and Scm3 from *K. lactis* and the functions of each domain and the relevant binding sites.

using these mutants that, with the exception of Cbf1, the localisation of kinetochore sub-complexes and the kinetochore assembly is dependent on CBF3 (Cai and Davis, 1990, Sorger et al., 1994, Poddar et al., 2004, Lang et al., 2018).This is reflected by the loss of kinetochore- microtubule interactions, when CBF3 function is lost (Sorger et al., 1994). The kinetochore dependency on CBF3 in budding yeast is likely due to the assembly of the kinetochore being dependent on Cse4 nucleosome deposition, which is facilitated by CBF3, hence why the CBF3 complex is not found in the assembled budding yeast kinetochore (Camahort et al., 2007, Stoler et al., 2007, Akiyoshi et al., 2010).

Later two additional functions have been attributed to the CBF3 complex. Mutations in *ndc10* shows phenotypic defects associated with spindle stability at anaphase and defective cytokinesis (Bouck and Bloom, 2005). In budding yeast, it has been shown that the spindle midzone hosts a number of mitosis associated proteins and this includes CBF3, where CBF3 has been transported to the midzone at anaphase, localised to the microtubule plus end (Bouck and Bloom, 2005). Its was proposed that the CBF3 complex may contribute to anaphase spindle stability, it may contribute to the pathway required for cytokinesis onset post chromosome segregation or the CBF3 complex initiates a pre-kinetochore at the microtubule plus end which then matures into the

full kinetochore complex as the cell cycle progresses (Bouck and Bloom, 2005). In addition, using *ndc10* mutants, it has been suggested that the CBF3 complex functions and regulates Septins and cytokinesis, in a function which is independent from its kinetochore functions (Gillis et al., 2005). Of the components of the kinetochore, only CBF3 subunit mutants alter septin organisation (Gillis et al., 2005). When CBF3 is inhibited at the G1 phase, Septin organisation is defective and when Septin dynamics are measured at anaphase, it was shown that CBF3 is needed for ring separation and disassembly of mother septin ring, and when defective, this contributed to aberrant polarised cell growth and cytokinesis (Gillis et al., 2005). This function is speculated to be linked to the Bir1 protein, which is a part of the chromosome passenger complex in budding yeast (Gillis et al., 2005).

Moreover, the CBF3 complex has been shown to alter the function of the SAC signalling (Gardner et al., 2001). The signalling activity has been mapped to Cep3, Ndc10 and Ctf13 (Gardner et al., 2001). When Cep3 is removed, checkpoint signalling is abolished (Gardner et al., 2001). When Ctf13 is removed, the cell cycle is delayed, due to spindle checkpoint signalling (Wang et al., 1993, Pangilinan and Spencer, 1996, Gardner et al., 2001). Dual mutants of Cep3 and Ctf13 showed that presence of both proteins is necessary for activation of the SAC and is required for maximum cell viability (Gardner et al., 2001). When Ndc10 is eliminated, the spindle checkpoint signalling is also removed (Tavormina and Burke, 1998). This influence on spindle checkpoint signalling is not present when other non-essential kinetochore subunits are removed or when Cse4 is removed (Gardner et al., 2001). It can be concluded that checkpoint signalling activity is dependent on all three of the subunits of the CBF3 complex (Gardner et al., 2001). Assuming that CBF3 would need to interact with other signal checkpoint proteins for its function, it has been speculated that the loss of CBF3 could indirectly alter interactions of the kinetochore with checkpoint signalling proteins (Gardner et al., 2001). Elimination of the individual subunits inhibits formation of the CBF3 complex, which in turn prevents recruitment of other essential kinetochore subunits including the Okp1 – Ame1 heterodimer and Mcm21 of the COMA complex and Mif2, and this in turn results in loss of spindle checkpoint signalling and the kinetochore. (Gardner et al., 2001). Evidence for an additional interaction between Skp1 and the spindle checkpoint Bub1 has been proposed and this interaction could also be involved in SAC signalling (Kitagawa et al., 2003).

1.5.2 Regulation of the CBF3 complex assembly

The assembly of the CBF3 complex is mediated by the chaperone Hsp90 and the co-chaperone Sgt1 (Kitagawa et al., 1999, Stemmann et al., 2002, Bansal et al., 2004, Lingelbach and Kaplan, 2004, Catlett and Kaplan, 2006, Bansal et al., 2009). Hsp90 has 1232 unique protein interactions, accounting for 20% of the yeast genome (Zhao et al., 2005, Breitkreutz et al., 2007).

Hsp90 functions downstream of the folding pathway where it is not required for *de novo* folding but it is involved in maintaining, activating or maturing proteins called clients (Nathan et al., 1997, McClellan et al., 2007). Budding yeast have two isoforms of Hsp90 called Hsc82, which is constitutively expressed and Hsp82, which is up regulated under stress (Borkovich et al., 1989). Due to the large selection of clients, which are unrelated in sequence, structure or function, Hsp90 is required to distinguish one client from the next (Prodromou et al., 1999, Panaretou et al., 2002, Richter et al., 2004, Zhao et al., 2005). This function is imparted by a class of proteins called co-chaperones (Sahasrabudhe et al., 2017). Sgt1 is a highly conserved co-chaperone with diverse cellular functions in eukaryotic organisms (Kitagawa et al., 1999, Dubacq et al., 2002, Lee et al., 2004, Lingelbach and Kaplan, 2004, Steensgaard et al., 2004, Catlett and Kaplan, 2006, da Silva Correia et al., 2007, Mayor et al., 2007, Zhang et al., 2008, Davies and Kaplan, 2010). In budding yeast it has been implicated in cell cycle regulation, adenylyl cyclase signalling and kinetochore assembly (Dubacq et al., 2002, Lingelbach and Kaplan, 2004, Catlett and Kaplan, 2006). In different eukaryotic organisms, Sgt1 typically interacts with different client proteins containing the Leucine Rich Repeat (LRR) motif, where the backbone of the motif is composed of predominantly leucine or isoleucine residues (Kobe and Kajava, 2001, Hoser et al., 2014, Hong and Hahn, 2016).

The need for the chaperone system in the context of the CBF3 complex has been extensively studied both genetically and biochemically (Kitagawa et al., 1999, Stemmann et al., 2002, Lingelbach and Kaplan, 2004, Catlett and Kaplan, 2006, Bansal et al., 2009). However the sequence of events, which lead to stability, and formation of the CBF3 complex is yet to be determined. The oligomeric state of Sgt1 has functional relevance to kinetochore assembly (Bansal et al., 2009). Sgt1 is dimeric both *in vitro* and *in vivo* and its been suggested that the serine/threonine kinase, CK2, negatively regulates dimerisation (Bansal et al., 2009). Phosphorylation of Serine 361 of Sgt1 inhibits dimerisation and results in non-viable budding yeast cells (Bansal et al., 2009). Hsp90 and Sgt1 need to be present to successfully express the intact CBF3 complex

(Kitagawa et al., 1999, Stemmann et al., 2002, Bansal et al., 2004, Lingelbach and Kaplan, 2004, Catlett and Kaplan, 2006, Bansal et al., 2009, Leber et al., 2018, Zhang et al., 2018). This is because the chaperone system is needed to stabilise and activate the Ctf13 protein using Skp1 (Bansal et al., 2004, Zhang et al., 2008). However, Sgt1 cannot be visualised or detected as part of the complex, suggesting that the function of the chaperones are transient (Bansal et al., 2004). Sgt1 as a co-chaperone provides the specificity to Hsp90 to bind Skp1 (Bansal et al., 2004, Catlett and Kaplan, 2006). As a result, the Sgt1 is able to bridge Hsp90 to Skp1, which is associated with Ctf13 (Catlett and Kaplan, 2006). This is reflected by studies which have shown Hsp90 cannot directly interact with Skp1, but Skp1 can interact with Sgt1 (Catlett and Kaplan, 2006). The transient nature of the interaction of Sgt1 with the CBF3 complex is highlighted by the low binding affinity of Sgt1 for Skp1 and Ctf13 (Bansal et al., 2004).

1.6 Structural studies of the CBF3 protein complex

The 'resolution revolution' in cryo electron microscopy (cryo-EM) has provided a plethora of structural data related to the CBF3 complex (Kühlbrandt, 2014, Leber et al., 2018, Zhang et al., 2018). Prior to this, structural studies were limited to truncated Cep3, Skp1 and truncated Ndc10, solved using X-ray crystallography (Schulman et al., 2000, Bellizzi et al., 2007, Purvis and Singleton, 2008, Cho and Harrison, 2012, Perriches and Singleton, 2012, Willhoft et al., 2017). Currently there are four CBF3 complex structures available including the CBF3 core complex, a truncated CBF3 core complex (CBF3ΔN), where the Cep3 N-terminal Zn_2Cys_6 cluster domain (Domain 1) was removed, and a truncated CBF3 complex where Cep3 Zn_2Cys_6 cluster domain was removed as well as domains 3 to 5 in Ndc10 (Fig 1.10). Recently an additional structure of the extended CBF3 complex bound to CEN3 DNA was published (Yan et al., 2018).

1.6.1 Cep3

Structure of Cep3 built from the cryo-EM density resembles the crystal structures solved previously (Bellizzi et al., 2007, Purvis and Singleton, 2008, Leber et al., 2018, Zhang et al., 2018). The Cep3 structure is predominantly composed of three domains, Zn_2Cys_6

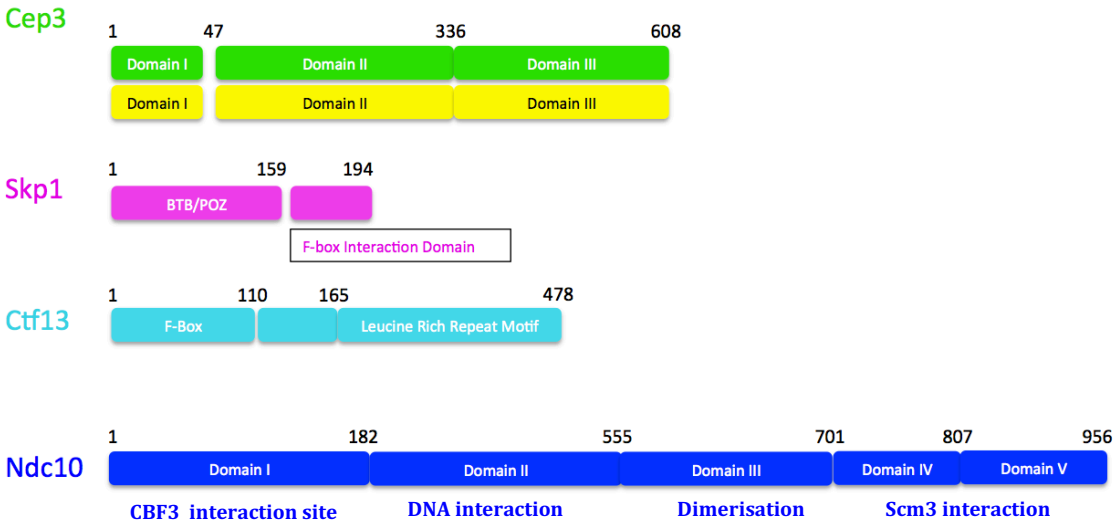


Fig. 1.10 **Domain organisation of CBF3 subunits.**(The diagram shows the name and function of specific domains of the individual subunits from the *S. cerevisiae* CBF3 complex.

cluster domain which shares strong homology to Gal4, followed by two predominantly α – helical domains (Fig 1.11.A) (Bellizzi et al., 2007, Purvis and Singleton, 2008). The third domain has a distinct solenoid structure composed of 13 short helices, which form 7 helical zigzags, which encircles the first helix of domain 2 next to the N-terminal domain (Bellizzi et al., 2007, Purvis and Singleton, 2008). The successive helices are oriented in an anti-parallel direction with the exception of the helices joined by two loops, one of which is located at the site of the C-terminal acidic patch domains (Bellizzi et al., 2007, Purvis and Singleton, 2008). The dimerisation interface, which enables the formation of the 2-fold axis of Cep3, spans the two domains (Bellizzi et al., 2007, Purvis and Singleton, 2008). The dimerisation is robust composed of mainly hydrogen bonds and side chain contacts, and the dimerisation is mediated differently to Gal4 like fungal transcription factor proteins, where typically the dimerisation is modulated by coiled coil domains (Hach et al., 2000, Bellizzi et al., 2007, Purvis and Singleton, 2008, Hong et al., 2008).

Unlike the crystal structures of Cep3, the cryo-EM density has Zn_2Cys_6 cluster domain present, but only one of the domains could be resolved (Fig 1.11.B) (Leber et al., 2018). The resolved domain was situated and stabilized by the heterodimer of Skp1 – Ctf13 (Leber et al., 2018). This domain displays pseudo-dyad symmetry and superimposition

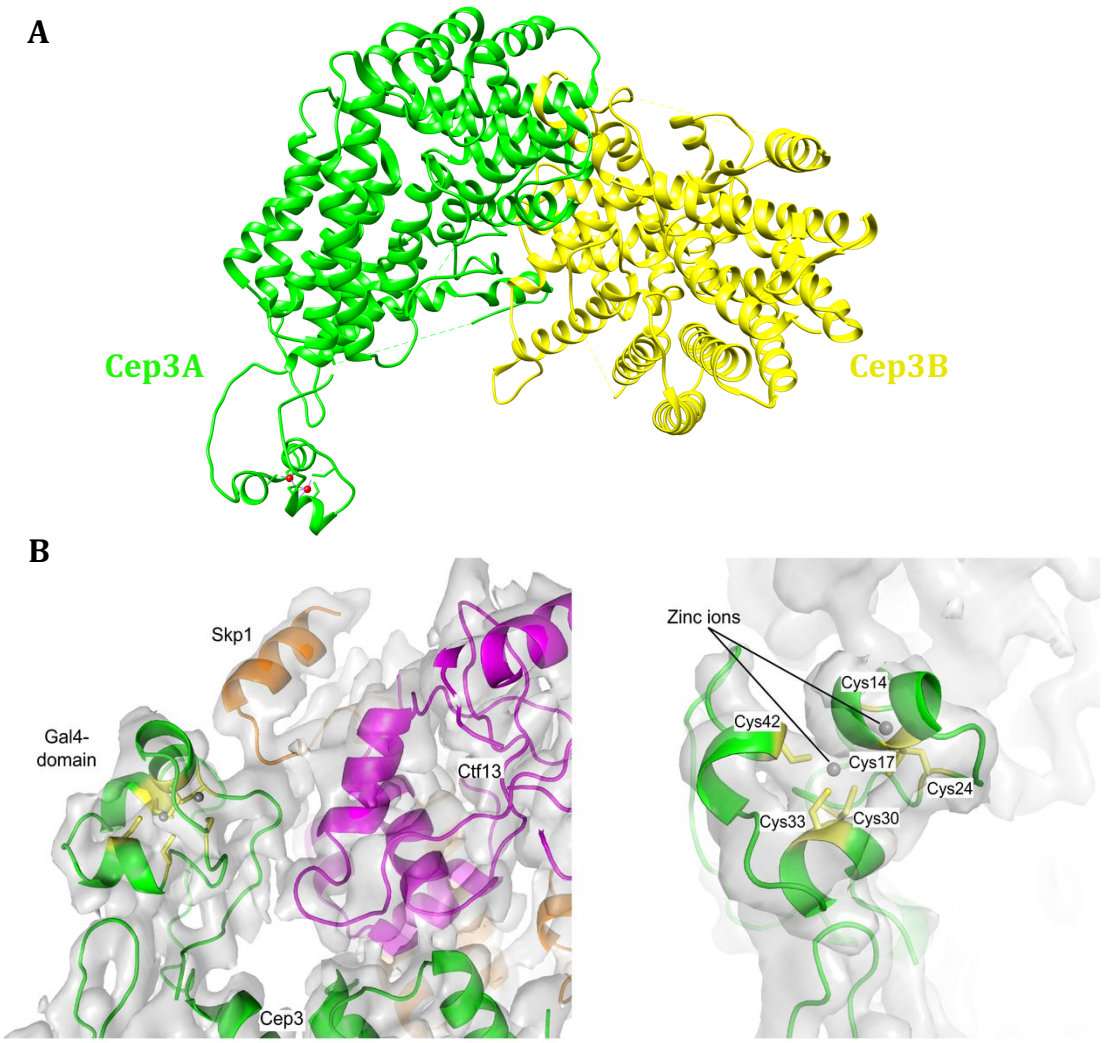


Fig. 1.11 **Structure of Cep3.**(A) Cep3 homodimer from 3.6 Å cryo-EM map of CBF3 core complex with the Zn_2Cys_6 domains, zinc highlighted in red. PDB ID: 6F07. (B) Closer view of the Gal4 domain docked at the Skp1-Ctf13 interface. This figure was taken from (Leber et al., 2018). EMDB ID: 4163.

with the crystal structure of a related Zn_2Cys_6 cluster domain from Hap1 bound to DNA, shows only one of the two orientations is consistent with the density (Leber et al., 2018). The second Zn_2Cys_6 cluster domain is not visible in the density, likely due to it being flexible and free for DNA binding (Leber et al., 2018). It's postulated the visible domain in the cryo-EM density is an inactive domain not capable of binding DNA and the release of the domain for DNA binding is linked to dephosphorylation of the protein Skp1 (Leber et al., 2018, Zhang et al., 2018).

The Zn_2Cys_6 cluster domain is expected to interact with two triplet sequences in CDEIII, CCG and TGT (Espelin et al., 1997, Purvis and Singleton, 2008). Mutations to CCG results in complete loss of centromere function, whilst mutations in TGT sequence shows that the sequence is important for fidelity of chromosome separation, but not essential (Jehn et al., 1991). This is reflected by DNA binding assays which show CCG is essential to Cep3 binding but TGT is not (Purvis and Singleton, 2008).

1.6.2 Skp1 and Ctf13

The crystal structures of Skp1 published previously corroborates with the cryo-EM structure (Schulman et al., 2000, Willhoft et al., 2017, Leber et al., 2018, Zhang et al., 2018). Skp1 has two domains, the larger BTB/POZ domain and the smaller F-box binding domain (Fig 1.12) (Schulman et al., 2000, Willhoft et al., 2017, Leber et al., 2018, Zhang et al., 2018). The N-terminal domain of Skp1 has a α/β structure where clusters of α-helices are flanked by β-sheets, and this fold is followed by a helical insertion and two helices from the C-terminal extension forming the F-box binding site ((Schulman et al., 2000, Willhoft et al., 2017, Leber et al., 2018, Zhang et al., 2018).

Previously, very little was known about the structure of Ctf13. This was attributed to the instability of Ctf13 resulting in the protein degrading quickly (Russell et al., 1999). To obtain a stable Ctf13, CBF3 binding partners and the chaperone system Hsp90 and Sgt1 are all required (Kitagawa et al., 1999, Stemmann et al., 2002, Lingelbach and Kaplan, 2004, Catlett and Kaplan, 2006, Bansal et al., 2009). The published cryo-EM structures of the CBF3 complex shows that the Ctf13 forms the central hub of the complex, interacting with Cep3, Skp1 and Ndc10 (Fig 1.16) (Leber et al., 2018, Zhang et al., 2018). From the structural analysis it can be deciphered that Ctf13 is composed of two distinct domains: an atypical F-box domain and a Leucine Rich Repeat (LRR) motif (Fig 1.12) (Leber et al., 2018, Zhang et al., 2018). The F-box domain is a α-helical domain joined to the LRR motif through a linker subdomain composed of 3-stranded anti-parallel β-sheet and a long α-helix that connects to the LRR (Leber et al., 2018, Zhang et al., 2018). The LRR motif is composed of 8 stranded parallel β-sheet surrounded by additional domains at each end, including a α - β subdomain at the C-terminal end of Ctf13 (Zhang et al., 2018). The C-terminal subdomain stabilises an acidic loop found on Skp1 to form a toroidal structure where the C-terminal subdomain of Ctf13 contacts the N-terminal F-box (Zhang et al., 2018).The toroid structure can be found in other F-box proteins such as TIR1 and COl1 plant hormone receptors (Tan

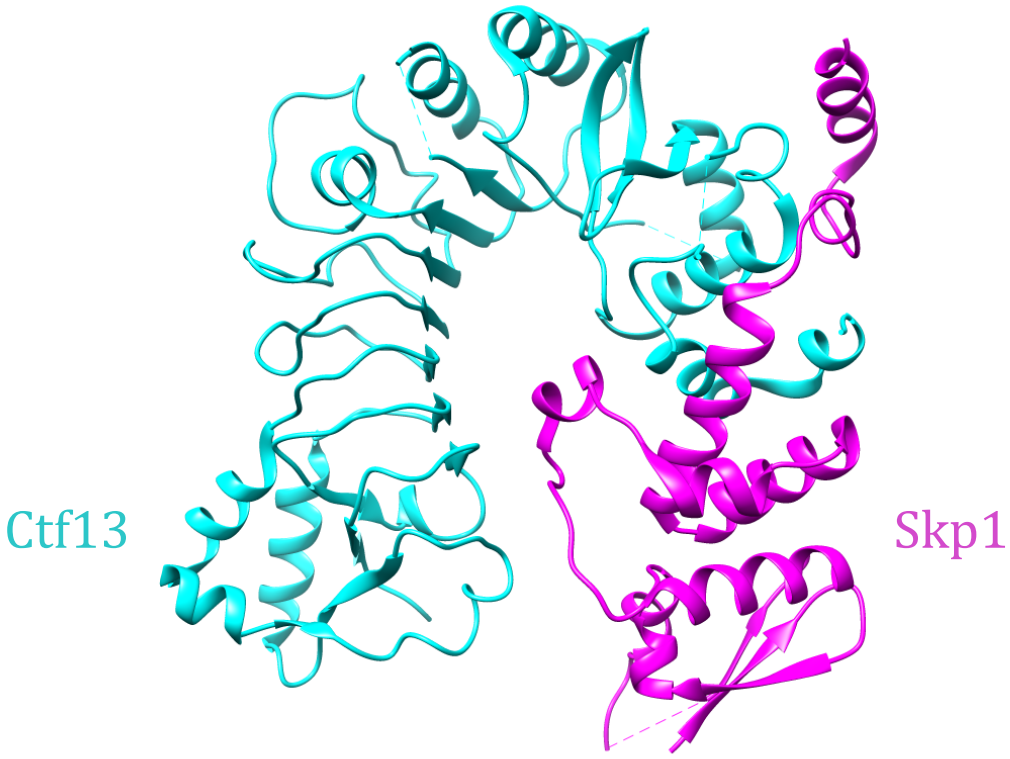


Fig. 1.12 **Structure of Skp1-Ctf13 heterodimer.**The atomic model of Ctf13 and Skp1 from the 4.2 Å truncated CBF3ΔN structure. PDB ID: 6GSA

et al., 2007, Sheard et al., 2010). Ctf13 however is distinct from these proteins because the toroid is formed by a combination of the LRR domain and the F-box domain, whereas in the receptor proteins, the toroid is formed only by the LRR domain (Zhang et al., 2018).

1.6.3 Ndc10

Limited tryptic proteolytic digestion of Ndc10 from *K. lactis* has shown that the protein is composed of 5 domains (Cho and Harrison, 2012). Combination of structural and biochemical analysis of Ndc10 from *S. cerevisiae* and *K. lactis* has revealed the first two domains: 1 and 2 are sufficient for interaction with CBF3 core complex (Fig 1.13) (Cho and Harrison, 2012, Zhang et al., 2018). Domain 2 is the primary DNA binding site but domain 1 and 3 also contribute to DNA binding affinity (Cho and Harrison, 2012, Perriches and Singleton, 2012, Zhang et al., 2018). Domain 3 also

contains the dimerisation interface (Cho and Harrison, 2012, Perriches and Singleton, 2012). The final two domains: 4 and 5, are responsible for association with Scm3 and the subsequent localisation of the Cse4 nucleosome to the point centromere (Cho and Harrison, 2012, Perriches and Singleton, 2012).

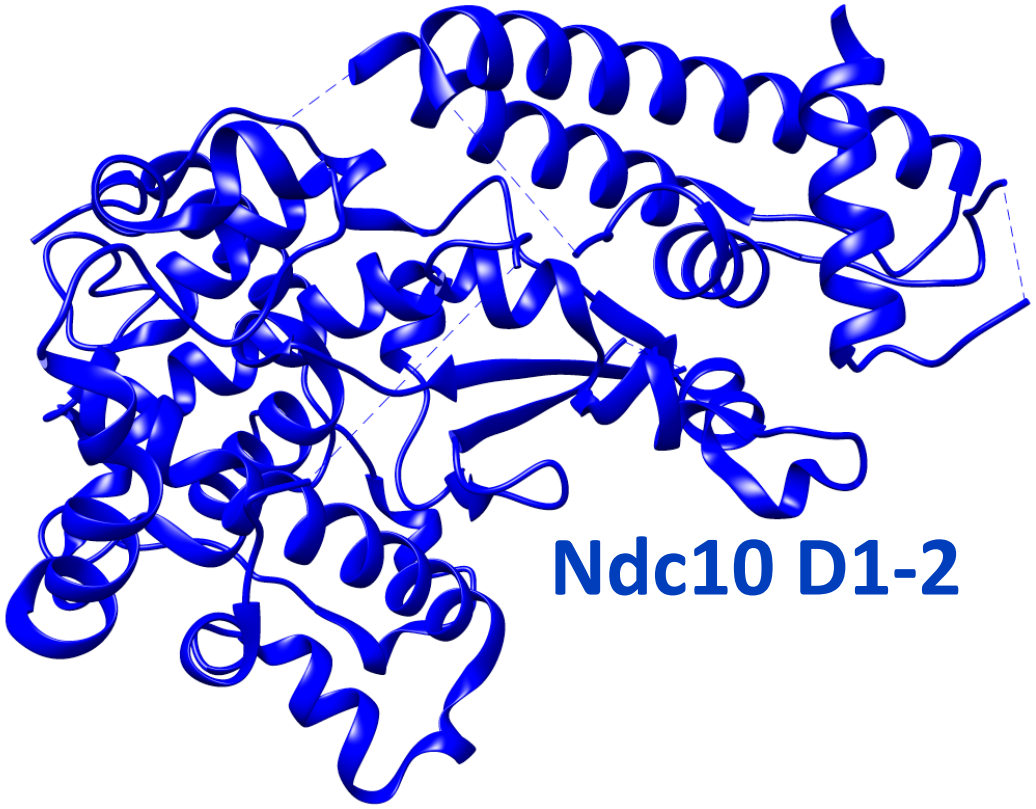


Fig. 1.13 **Structural studies of Ndc10.**(A) 1.9 Å crystal structure of Ndc10 DNA binding region (Domains 1 and 2) from *S. cerevisiae*. PDB ID: 4ACO.

Both x-ray crystallography and the later cryo-EM data for Ndc10 is limited to domain 1 and 2 (Cho and Harrison, 2012, Perriches and Singleton, 2012, Zhang et al., 2018).The two domains can be divided into two α-helical lobes surrounding a central β-sheet (Fig 1.13) (Cho and Harrison, 2012, Perriches and Singleton, 2012). The N-terminal lobe consists of an anti-parallel four-helix bundle whilst the C-terminal lobe is composed of several long loops interspersed with short helices (Cho and Harrison, 2012, Perriches and Singleton, 2012). Analysis of the electrostatic surface show, the curved surface is composed of residues, which are predominantly positively charged, and has the highest number of conserved residues in comparison to the rest of the protein regions (Cho and Harrison, 2012, Perriches and Singleton, 2012).

1.6.4 CBF3 core complex architecture

The CBF3 core complex composed of Cep3, Skp1 and Ctf13 forms a horse-shoe –like structure with Cep3 homodimer forming one end and the Skp1 – Ctf13 heterodimer joining a Cep3 monomer at the outside edge to complete the structure (Fig 1.14) (Leber et al., 2018, Zhang et al., 2018). Only one heterodimer appears to be present and attempts to dock a second heterodimer resulted in steric clash (Leber et al., 2018). The distal Cep3 monomer helps complete the horseshoe shape (Leber et al., 2018, Zhang et al., 2018).

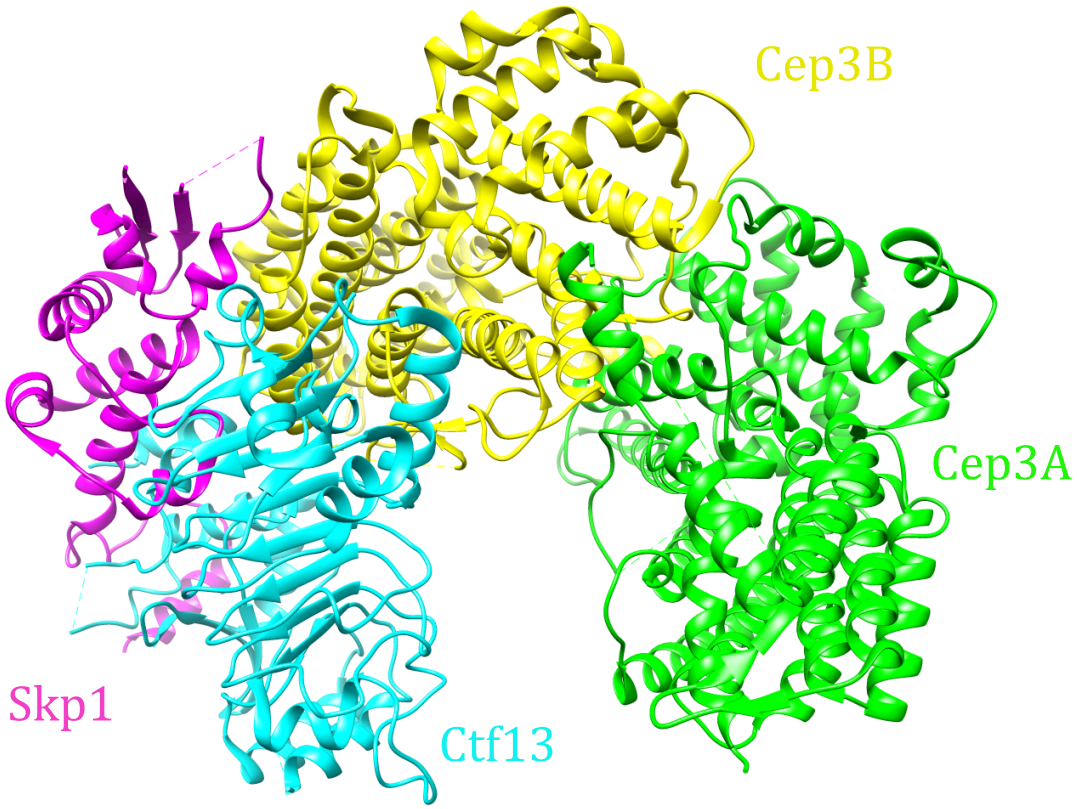


Fig. 1.14 **Structure of the CBF3 Core Complex.**3D model of the CBF3 core complex composed of Cep3 homodimer, and Ctf13 – Skp1 heterodimer. The N-terminal Zn_2Cys_6 domain of Cep3 is truncated in the structure. PDB: 6FE8.

A notable feature of the CBF3 core complex, is the channel formed by the three proteins (Fig 1.14) (Leber et al., 2018, Zhang et al., 2018). The central channel is approx. 30 Å in diameter and is wide enough to accommodate a DNA duplex (Zhang et al., 2018). The channel is lined with basic residues from Cep3 and Ctf13 which are

strongly conserved between Saccharomycetaceae (Zhang et al., 2018). Ctf13 contributes a number or arginines and lysines which extend into the channel from the inter LRR turns of LRR 1 to 6 (Fig 1.15) (Zhang et al., 2018). Each of the Cep3 protomers also contribute lysine and arginine to the channel (Zhang et al., 2018). When these key Cep3 and Ctf13 basic residues are mutated, there is a noticeable loss in DNA binding, as shown through EMSA assays ((Zhang et al., 2018).

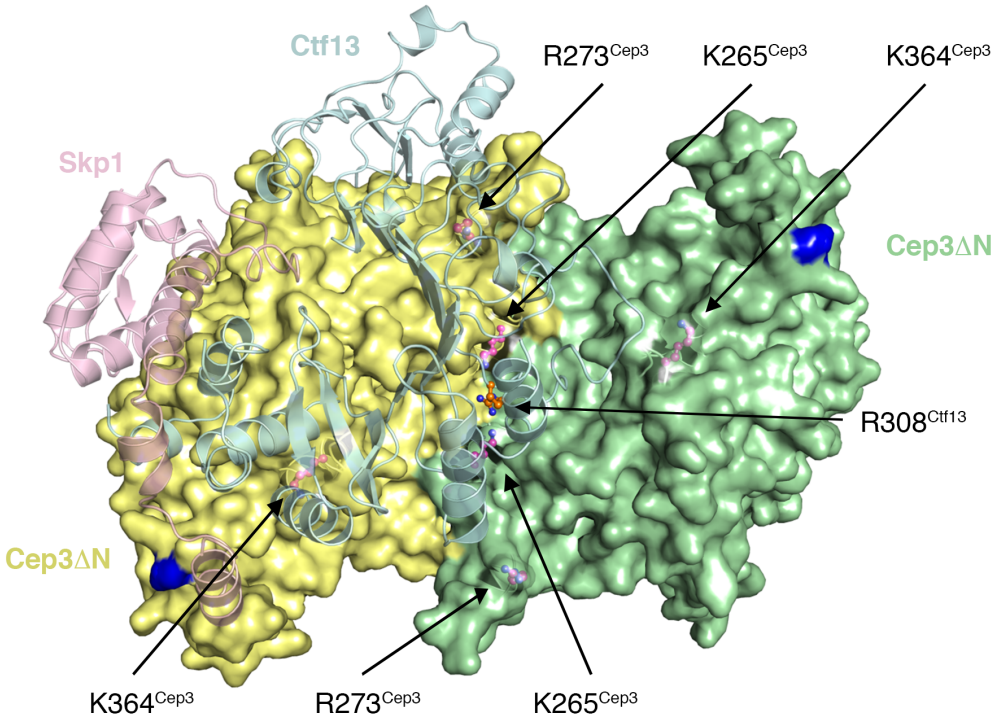


Fig. 1.15 **DNA binding channel of the CBF3 core complex.**Close up view of the potential DNA binding channel, lined with arginine and lysine residues from Cep3 and Ctf13. This figure was taken from (Zhang et al., 2018).

1.6.5 CBF3 complex with Ndc10

Cryo-EM map of the CBF3 core complex with Ndc10 domains 1 and 2 (Fig 1.16) without centromeric DNA revealed flexibility in Ndc10 relative to the rest of the complex (Zhang et al., 2018). Docking of the crystal structure of Ndc10 Domains 1 and 2 into the cryo-EM density map shows structural elements at the interface not present

in the crystal structure (Perriches and Singleton, 2012, Zhang et al., 2018). Ndc10 domains 1 and 2 are sufficient for association with the CBF3 core complex, with Ndc10 forming interaction with Ctf13 and Skp1 across a small interface (Zhang et al., 2018). The interaction is facilitated by a stretch of conserved residues at the N-terminus of Ctf13 and C-terminal residues of Skp1 (Zhang et al., 2018). The cryo-EM structure shows Ndc10 forming a hinge with the core complex, which imparts both flexibility to Ndc10 and when combined with the small interaction interface, it explains why the Ndc10 readily separates from the core CBF3 complex during protein purification (Leber et al., 2018, Zhang et al., 2018).

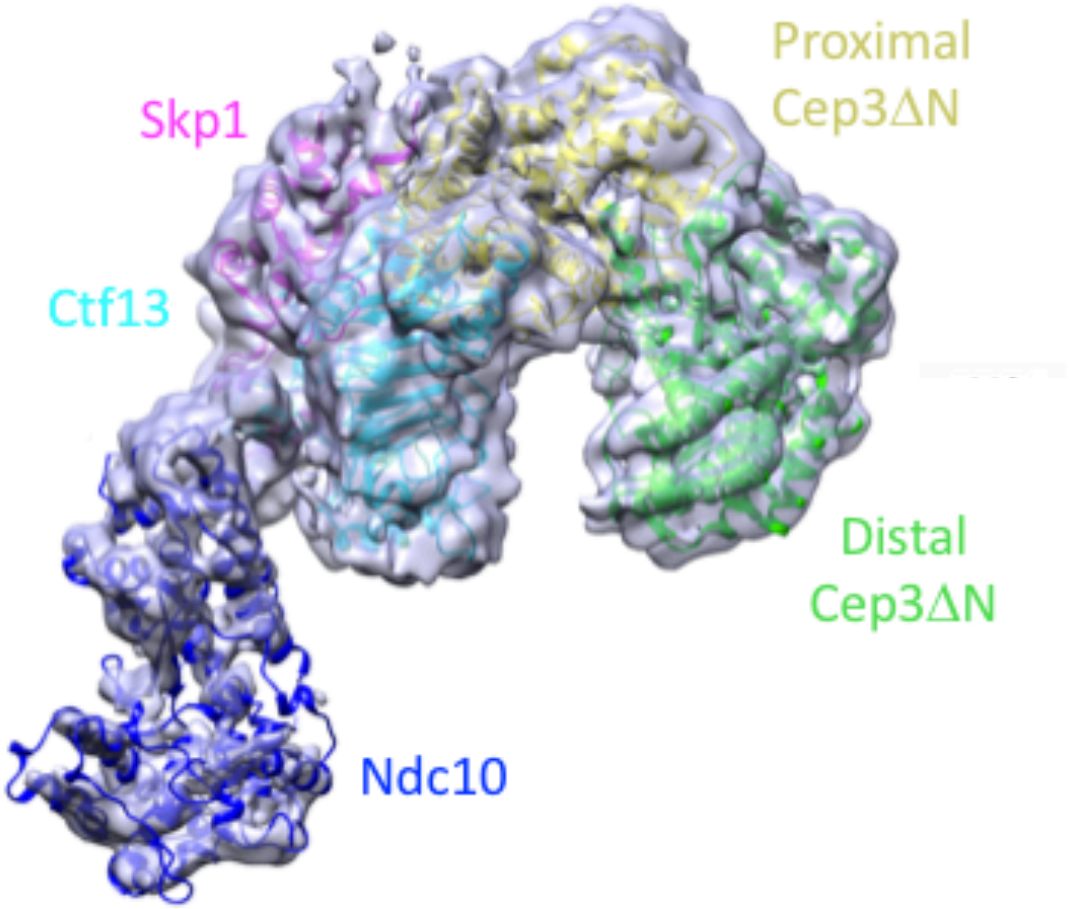


Fig. 1.16 **Structure of CBF3 core complex with Ndc10 D1-2.** 4.4 Å cryo-EM map of the CBF3 complex with the DNA binding region of Ndc10. This figure was taken from (Zhang et al., 2018).

Towards the end of the project a 4.4 Å cryo-EM structure of the CBF3 complex bound to 147 bp CEN3 DNA (Fig 1.17) was published (Yan et al., 2018). The structural analysis conducted, identified a dimeric CBF3 complex, composed of two CBF3 core complexes bridged by a dimeric Ndc10 (Yan et al., 2018). The structure of the core CBF3 complexes was identical to the two structures of the core CBF3 complexes published previously (Leber et al., 2018, Zhang et al., 2018). The density of the Ndc10 also resembled the density of the previously resolved Ndc10 domains 1 to 2 bound to the CBF3 core complex structure (Zhang et al., 2018). In the CBF3-CEN3 map, Nd10 only interacts with Ctf13, via Ndc10 domains 1 and 2, across a small interface, similar to the CBF3 core complex bound to Ndc10 domains 1 and 2 map (Zhang et al., 2018). The map also shows Ndc10 domains 1 and 2 have dimerisation activity when engaged with the CEN3 DNA (Fig 1.17) (Yan et al., 2018). The self-association occurred through a small interface and is expected to be weak (Yan et al., 2018). Biophysical studies of Ndc10 domains 1 and 2 show that the protein is monomeric in solution without DNA, which suggests that the association of two Ndc10 DNA binding domains could be induced by the DNA (Cho and Harrison, 2012, Perriches and Singleton, 2012, Yan et al., 2018).

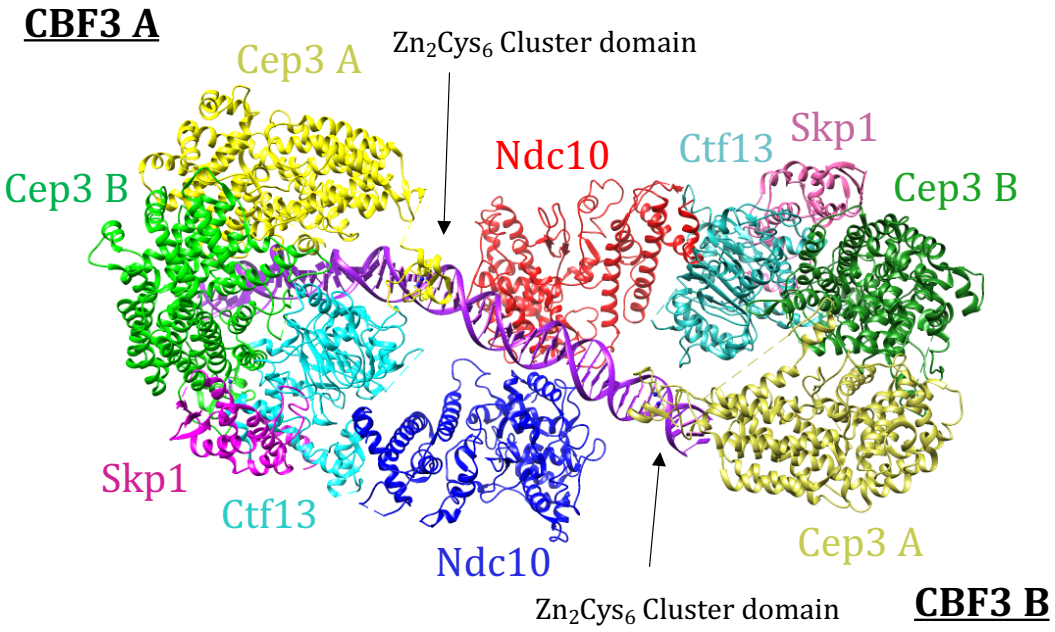


Fig. 1.17 **Structural model of the CBF3 complex bound to 147 bp CEN3 DNA.** Model of the dimeric CBF3 bound to the CEN3 DNA. The individual subunits are labelled and the arrow show the position of the Zn_2Cys_6. PDB ID: 6GYS

1.6.6 Proposed DNA Binding Models

At present, there are two proposed models for how the CBF3 complex binds to centromeric DNA. The first model (Fig 1.17) was proposed in the recently published CBF3-CEN3 study where the centromeric DNA induces CBF3 dimerisation (Yan et al., 2018). The dimeric CBF3 complex binds to a gently bent CEN3 DNA duplex, occupies 6 bp of CDEII, the full CDEIII sequence and 24 bp extending from CDEIII, protecting 56 bp of CEN DNA (Lechner and Carbon, 1991, Yan et al., 2018). The two core CBF3 complexes, CBF3 A and CBF3 B bind to different regions of the DNA duplex (Yan et al., 2018). CBF3 A engages with DNA using both Cep3 and Ctf13 (Yan et al., 2018). The Zn_2Cys_6 domain of Cep3, situated next to the heterodimer Ctf13-Skp1, interacts with the heterodimer and is stabilised as observed in the CBF3 core complex structure (Leber et al., 2018, Yan et al., 2018). The Zn_2Cys_6 of the other Cep3 protomer engages with the CCG motif (Yan et al., 2018). The Cep3 makes additional contact with CEN DNA via an α-helix with the TGT motif (Yan et al., 2018). In between the CCG and the TGT contacts, a Ctf13 loop inserts into the minor groove to interact with the phosphate backbone of the DNA, as was shown by the Protein-DNA crosslinking experiments (Espelin et al., 1997, Yan et al., 2018). CBF3 B engages further upstream of CBF3 A, engaging with CEN3 DNA using the Zn_2Cys_6 domain of the Cep3 promoter, 20 bp away from the CCG motif, interacting with a new DNA sequence yet to be characterised (Yan et al., 2018).

The misalignment of the CBF3-CEN3 model with previously published data is a major limitation of this study. Genetic and biochemical studies have established that the CBF3 complex was approximately 445 kDa in size which matches a monomeric CBF3 and a combination of gel filtration and glycerol velocity gradients determined that the CBF3 complex contained a homodimer of Cep3, heterodimer of Ctf13-Skp1 and a homodimer of Ndc10 (Lechner and Carbon, 1991, Espelin et al., 1997, Russell et al., 1999). The DNA-Protein cross-linking studies proposed two forms of the CBF3 complex: a monomeric CBF3 complex and an extended complex which is composed of the monomeric complex bound by an additional homodimer of Ndc10 (Espelin et al., 1997). These calculated molecular mass of the protein complex and the calculated stoichiometry from all of these previous studies does not align with the conformation observed in the CBF3-CEN3 study (Yan et al., 2018). The presence of a dimeric CBF3 complex raises questions regarding the Cse4 nucleosome recruitment. The questions include: does the dimeric conformation of the CBF3 change the number of

Cse4 nucleosomes recruited to the point centromere? how does the point centromere accommodate a dimeric CBF3 complex, Cse4 nucleosome and two copies of the CTF19 complex? Therefore, in order for this model to become the accepted mode of CBF3 binding, the data requires further experimental validation to ensure the results were not a consequence of the experimental setup.

The second model (Fig 1.18) is built using a combination of the cryo-EM map of the CBF3 core complex with Ndc10 DNA binding domain and the high resolution crystal structures of Ndc10 DNA binding domain to map the DNA pathway in the CBF3 complex (Cho and Harrison, 2012, Perriches and Singleton, 2012, Zhang et al., 2018). The model proposes that the monomeric CBF3 complex binds to CDEIII, promotes curvature of CDEII to facilitate Cse4 nucleosome recruitment (Zhang et al., 2018). The Ndc10 DNA binding site is perpendicular to the central channel and the DNA must bend significantly to satisfy both binding sites (Zhang et al., 2018). The CCG half site of CDEIII was located at the narrow end of the channel, next to Ctf13 and Cep3 whilst the TGT is located in a different environment at the solvent accessible end of the channel, which is more open (Zhang et al., 2018). The model extends 21-30 base pairs downstream of CCG sequence which means that the CDEIII sequence is extended to approximately 50 base pairs, consistent with the DNA foot print experiments which found CBF3 protecting approximately 56 base pairs (Lechner and Carbon, 1991, Zhang et al., 2018). The position of the second Ndc10 protomer is ambiguous in this model but the curvature of the C-terminus of Ndc10 DNA binding domain suggests that the second Ndc10 protomer is likely oriented towards CDEII (Zhang et al., 2018). The model accommodates both right-handed and left-handed DNA supercoil, but the model provides greater weight towards a right-handed supercoil because in that conformation the CDEI would be in close proximity to Ndc10 domain 1, consistent with CDEI bound Cbf1 interacting with Ndc10 domain 1 (Cho and Harrison, 2012, Zhang et al., 2018).

The second CBF3 binding model is consistent with features associated with CBF3-CDEIII DNA binding which were identified previously, unlike the first model (Lechner and Carbon, 1991, Espelin et al., 1997, Russell et al., 1999). However, the second model is limited by the experimental design and the method of modelling is prone to error as a result. The first limitation relates to the Ndc10. The determined structure used to model the DNA pathway contains only the DNA binding domain of Ndc10, with the dimerisation domain absent. This means information about the second Ndc10 protomer is missing, including the position of the second Ndc10 protomer, how the presence of two Ndc10 binding sites and the channel alters the DNA pathway and how

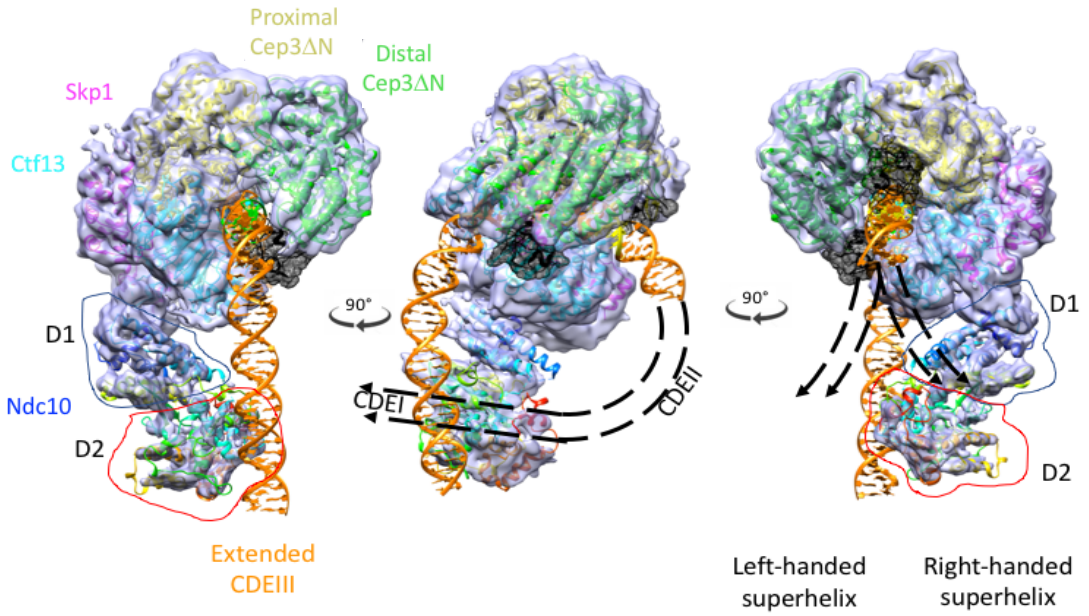


Fig. 1.18 **Monomeric CBF3-CDEIII binding model.** Model of the monomeric CBF3 bound to the CDEIII. The individual subuunits are labelled and D1 and D2 represent domain 1 and 2 of Ndc10. The dotted arrow represents the CDEI and CDEII sequences and their projected direction and orientation. This figure was taken from (Zhang et al., 2018).

the presence of two Ndc10s changes the overall architecture of the complex. Secondly, the determined structure does not contain the Zn_2Cys_6 cluster domain of Cep3. Whilst the model utilises the fungal transcription factor Gal4 Zn_2Cys_6 cluster domain which has sequence conservation with Cep3, data is needed to show that the Cep3 domain adopts the same conformation and is positioned in the way that the model assumes. The final limitation relates to the determined structure not having DNA bound. The model assumes that the CBF3 complex binds to the centromeric DNA via the channel and the Ndc10 DNA binding domain. Without the DNA present, its unknown exactly how the complex binds the DNA and whether DNA binding induces changes to the CBF3 architecture. Similarly to the first binding model, this proposed mechanism also requires further experimental validation.

1.6.7 Phosphoregulation of CBF3-CDEIII binding activity

The role of phosphorylation in CBF3 – CEN DNA binding has also been a topic of discussion in the field. Initial studies on the CBF3 protein complex determined that CBF3 DNA binding activity was related to the phosphorylation of one of CBF3 subunits (Lechner and Carbon, 1991). When the complex was completely dephosphorylated there was a distinct loss of centromere DNA binding activity of CBF3 (Lechner and Carbon, 1991). Later the role of phosphorylation and CBF3 activity was attributed to Skp1-mediated activation of the unstable Ctf13 (Kaplan et al., 1997). Phosphorylation of Skp1 was proposed to be required to activate Ctf13 by Skp1 to then form the core CBF3 complex (Kaplan et al., 1997). In the absence of the phosphorylation activity, Ctf13 is targeted for proteosomal degradation (Kaplan et al., 1997). This sugested the phosphorylation and ubiquitination pathways work in synergy to regulate the level of the CBF3 complex (Kaplan et al., 1997). A subsequent study disputed this hypothesis, by showing complete dephosphorylation of Ctf13 and Skp1 has no affect on the CBF3 protein complex formation and CBF3 complex – centromere binding activities (Stemmann et al., 2002). It was proposed that because Ctf13 can form a complex with Ndc10 and Cep3 when expressed in insect cells without Skp1, that the phosphorylation on Ctf13 can substitute for and displace Skp1 for complex formation or the Ctf13 phosphorylation in insect cells can stabilise the protein, removing the need for Skp1- Sgt1-Hsp90 function (Stemmann et al., 2002).

A recent study has linked CBF3 binding activity to the phosphorylation status of Skp1 (Leber et al., 2018). The cryo-EM structure showed the Gal4 like domain in inactive conformation, situated at the Skp1 and Ctf13 interface (Figure 1.17.C). Using electrophoretic mobility shift assays (EMSA), it was shown that in order for the CBF3 core complex to bind centromere DNA sequence, Skp1 needs to be dephosphorylated (Leber et al., 2018). This meant a candidate protein phosphatase activates the core complex, which is then able to bind centromere DNA and conduct its functions (Leber et al., 2018).

1.7 Project aims

The CBF3 complex forms the central hub of the kinetochore and is a pre-requisite for the recruitment of the other kinetochore components and the specialised centromeric

nucleosomes(Sorger et al., 1994, Camahort et al., 2007, Pagliuca et al., 2009, Lang et al., 2018).In the absence of the CBF3 complex, the outer kinetochore proteins are no longer able to attach to the microtubules (Sorger et al., 1994). The focus has now shifted towards building a mechanistic understanding of the protein complex to determine how the complex assembles, how the complex recognises and binds to centromeric DNA and how the association of the complex with centromeric DNA facilitates the recruitment of the Cse4 nucleosome and the CTF19 complex in order to initiate a set of events which facilitate chromosome segregation.

To build a mechanistic understanding of CBF3 function, it is important to decipher how the protein complex is able to recognise and bind specifically to CDEIII. Structural studies of the CBF3 core complex have identified a central channel, lined with basic residues, which is sufficiently wide to accommodate a DNA duplex (Zhang et al., 2008). This is supported by the CDEIII-CBF3 cross linking study which identified interactions between CDEIII and both Cep3 and Ctf13 (Espelin et al., 1997). These data suggested a model in which the DNA binds to the channel. An additional CBF3 structure, with the Zn_2Cys_6 cluster domains present was also published (Leber et al., 2018). Only one Zn_2Cys_6 cluster domain was visible in the EM map and it was speculated that the domain adopted an inactive conformation (Leber et al., 2018). The publication also presents a different CBF3-CDEIII binding model where the CDEIII DNA lies across the core CBF3 complex and does not thread through the channel, but the paper does suggest the channel could serve as a potential DNA binding site (Leber et al., 2018). Sequence homology suggests that the Zn_2Cys_6 cluster domain is a DNA binding domain and DNA-protein cross linking studies identified interactions between Cep3 and two triplet sequences, CCG and TGT studies (Lechner, 1994, Schjerling and Holmberg, 1996, Espelin et al., 1997, Purvis and Singleton, 2008). The CCG sequence is essential for both accurate chromosome segregation and CDEIII binding specificity , which is not the case for TGT (Jehn et al., 1991, Purvis and Singleton, 2008, Leber et al., 2018). The recent CBF3 core complex study suggests only the CCG triplet serves as Cep3 binding site (Leber et al., 2018).

Therefore, two questions arise from these early structural studies: where is the CDEIII binding site on the core complex and do the Zn_2Cys_6 cluster domains provide a canonical binding mode? To address these questions the first aim of my PhD project was to purify the CBF3 core complex bound to CDEIII and determine its structure using EM and conduct EMSA assays on the unbound CBF3 core complex.

Recent studies have shown that the CBF3 core complex ability to bind centromeric DNA is phosphoregulated and as a result the protein complex must undergo phosphatase treatment to bind CDEIII DNA (Leber et al., 2018). This contradicts earlier studies which proposed that one or more of the CBF3 subunits must be phosphorylated to achieve centromeric DNA binding and work from our laboratory in which binding to a short CDEIII oligonucleotide is routinely observed in a complex prepared using phosphatase inhibitors (Lechner and Carbon, 1991, Zhang et al., 2018). Further investigation is required in order to begin to decipher the complexity of phosphoregulation in CBF3-CDEIII binding. Therefore, the second aim of my PhD project was to determine the effects of dephosphorylation of the CBF3 complex on CDEIII binding, using EMSA assays.

The orientation of Ndc10 relative to the CBF3 core complex, positioned at one end of the central channel, led to a model in which the DNA adopts a non-linear conformation in order to satisfy binding to both the channel and the Ndc10 DNA binding domains which are distant from the channel (Zhang et al., 2018). The CBF3 complex in this study also lacks the dimerisation domain of Ndc10 and consequently the second protomer of the Ndc10 dimer, therefore the influence of the rest of the Ndc10 polypeptide and a second DNA binding domain on CDEIII binding is not accounted for in the model and remains unknown (Zhang et al., 2018). The third aim of my PhD project was to decipher how the presence of a Ndc10 dimer would change the architecture of the CBF3 complex, in comparison to the CBF3 core complex structure with a monomeric Ndc10 DNA binding domain and determine the resulting implications for centromeric DNA binding. To address this my goal was to assemble and purify the full CBF3 complex (including the Ndc10 dimer) in the presence and absence of CDEIII DNA and gain structural information using EM.

Towards the end of the project, a structure of the CBF3 complex bound to CEN3 DNA was published, showing a dimeric CBF3 complex bound to CDEIII (Yan et al., 2018). The dimeric complex consists of two copies of CBF3 core complex connected by two Ndc10 molecules with a stoichiometry of 4:2:2:2 (Cep3:Ndc10: Ctf13:Skp1) (Yan et al., 2018). The genetic and biochemical studies of the CBF3 complex all proposed a protein complex that had a molecular weight of approximately 445 kDa, which corresponds to a monomeric CBF3: a single CBF3 core complex and two copies of Ndc10 (Lechner and Carbon, 1991, Espelin et al., 1997, Russell et al., 1999, Zhang et al., 2018, Leber et al., 2018). The hydrodynamic data obtained from gel filtration and glycerol velocity gradients determined that the CBF3 complex was composed of

a Cep3 homodimer, Ndc10 homodimer and a heterodimeric Ctf13 and Skp1 (Russell et al., 1999). As a result, two different models of the CBF3 complex architecture and DNA binding models have now been proposed an further study is required to validate which is correct (Zhang et al., 2018, Yan et al., 2018).

Three questions arise as a result of this new study: does the CBF3 adopt a monomeric or dimeric structure, does the CBF3 architecture differ before and after centromeric binding and is the dimeric conformation of CBF3 a consequence of the experimental method or can it be replicated in an alternative study? To address these questions the final aim of this PhD project was to perform comparison study of the determined CBF3 complex structure with and without CDEIII bound in this thesis to the published CBF3-CEN3 structure, to identify similarities, differences and confirm whether this study validates or disputes those findings.

Chapter 2

Materials and Methods

2.1 Introduction

This chapter details the general protocols and the materials used to conduct the molecular biology, protein preparation and the subsequent structural and biochemical experiments used to characterise the CBF3 complex and the individual subunits.

2.2 General protocols

2.2.1 Media preparations

Luria-Bertani (LB) and Terrific Broth (TB) media (Merck) were prepared to an appropriate volume by suspending 25 g of LB powder and 47.6 g of TB powder in 1 litre of distilled water. 2x Tryptone Yeast Extract (TY) broth was prepared using 1.6 % Tryptone, 1% Yeast extract, 0.5% NaCl

Non-selective yeast extract peptone dextrose (YEPD) media was prepared by suspending 15 g of YPD powder (Fisher Scientific), 0.0165 g of adenine sulphate (Formedium), and 0.0165 g of L- tyrosine (Formedium) in 300 mL of distilled water with 100 μg/mL Ampicillin. Selective YEPD media was prepared by suspending 2.07 g of yeast nitrogen base without amino acids (Formedium), 0.0165 g of complete supplement mixture

without tryptophan and uracil, 0.0165 g of adenine sulphate (Formedium) and 0.0165 g of L- tyrosine (Formedium) in 300 mL of distilled water with 100 μg/mL Ampicillin and 2% glucose. The selective YEPD expression media does not contain tryptophan or uracil.

2.2.2 Agar plates

LB agar plates were prepared by suspending 40 g of LB agar powder (Fisher Miller Reagents) per litre of distilled water with the appropriate antibiotic (100 μg/mL Ampicillin, 50 μg/mL Kanamycin and 34 μg/mL Chloramphenicol) added prior to plating.

Non-selective YEPD agar plates were prepared by suspending 15 g of YPD powder (Fisher Scientific), 0.0165 g of adenine sulphate (Formedium), 0.0165 g of L- tyrosine (Formedium) and 6.6 g of agar (Formedium) in 300 mL of distilled water with 100 μg/mL Ampicillin. The selective YEPD agar does not contain tryptophan or uracil.

Selective YEPD agar plates were prepared by suspending 2.07 g of yeast nitrogen base without amino acids (Formedium), 0.0165 g of complete supplement mixture without tryptophan and uracil, 0.0165 g of adenine sulphate (Formedium), 0.0165 g of L- tyrosine (Formedium) and 6.6 g of agar (Formedium) in 300 mL of distilled water with 100 μg/mL Ampicillin and 2% glucose.

2.2.3 DNA gel electrophoresis

0.5 – 1.5% agarose gels (Fisher Miller Reagents) were prepared using either Tris, Boric acid and EDTA (TBE) buffer (45 mM Tris-borate, 1 mM EDTA) or Tris, Acetic acid and EDTA (TAE) buffer (40 mM Tris-acetate and 1 mM EDTA). The Gel was inserted into the Mini-Sub® Cell GT System (BIORAD) and the electrophoresis was conducted at 100 V for 50 minutes, unless otherwise stated.

2.2.4 DNA sequencing

DNA sequencing was used to check the fidelity of the site directed mutagenesis reactions. The sequencing reactions were shipped to GATC (Eurofins) who conducted the analysis. The results were analysed using ApE plasmid editor.

2.2.5 DNA annealing

Each oligonucleotide was prepared and stored at 100 μM stock concentration, in buffer (15 mM Tris, 50 mM NaCl, 1 mM $MgCl_2$). Equal volume of each oligonucleotide was added together and transferred to a SureCycler 8800 Thermal Cycler (Agilent). The sample was heated to 98°C for 10 minutes and then over 77 cycles, the temperature was decreased by 10°C every minute. At 20°C, the sample was removed. The oligonucleotides used are listed in Table 2.15.

2.3 Molecular biology methods

2.3.1 Plasmid propagation, purification and quantification

Plasmid stock was prepared by inoculating a colony into 5 mL of LB media with the appropriate antibiotic and incubated for 12-16 hours at 37°C. The cell culture was centrifuged at 3,214 x g at 4°C to obtain the cell pellet. The plasmid was purified using the QIAprep Spin Miniprep Kit (QIAGEN) as instructed. Concentration of the purified DNA was calculated using NanoDrop® Spectrophotometer NDC-1000.

2.3.2 Maintenance and preparation of competent cells

The competent cells prepared and maintained during this project are listed in the table below (Table 2.1).

Preparation of chemically competent Top10 *E. coli* cells
Aliquot of Top10 *E. coli* cells was streaked onto an LB agar plate with no antibiotic. The plate was incubated at 37°C for 12-16 hours. A single colony was inoculated in to

Cell strain	Organism	Purpose
One Shot^{TM} chemically competent cells	*E. coli*	Plasmid propagation
BL21-Codon Plus (DE3)-Ril chemically competent cells	*E. coli*	Protein expression
ElectroMAX^{TM} DH5-α electrocompetent cells	*E. coli*	Yeast expression cassette preparation
TG1 Electrocompetent cells	*E. coli*	Yeast expression cassette preparation
BCY123	*S. cerevisiae*	Yeast protein expression

Table 2.1 **List of Cell strains.** List of the different cell strains used to conduct different experiments.

1 mL of LB medium with no antibiotic and incubated at 37°C for 12-16 hours. The 1 mL of culture was inoculated in to 100 mL of LB medium with 20 mM $MgSO_4$ and incubated at 37°C until OD_{600nm} of 0.6 is reached. The cells were harvested at 3,214 x g at 4°C for 5 minutes. The cell pellet was resuspended in 40 mL TFB1 buffer (30 mM Potassium Acetate, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 100 mM RbCl, 15% glycerol) at pH 5.8 and incubated on ice for 5 minutes. The cells were pelleted at 3,214 x g at 4°C for 5 minutes. The cell pellet was resuspended in 4 mL of TFB2 buffer (100 mM MOPS, 75 mM $CaCl_2$, 10 mM RbCl, 15% Glycerol) at pH 6.5. The cells were incubated on ice for 60 minutes, aliquoted and flash frozen at -80°C. This protocol was adapted from (Hanahan, 1983).

Preparation of chemically competent *BCY123 S. cerevisiae* cells

The cells were streaked on to a non-selective YEPD agar and incubated at 30°C for 3 days. A single colony was inoculated in to 6 mL of non-selective YEPD medium and incubated at 30°C for 12-16 hours. 1 mL of the culture was then inoculated in to 40 mL of non-selective YEPD medium and incubated at 30°C, 220 rpm until OD_{600nm} of 0.6. The cells were harvested by centrifugation at 1,316 x g, 4°C for 5 minutes. The pellet was resuspended into 40 mL of sterile filtered water and centrifuged at 1,316 x g, 4°C for 5 minutes to wash the pellet. Next the pellet was resuspended in 40 mL of Lithium Acetate Buffer (100 mM Lithium acetate dihydrate, 10 mM Tris pH 7.4, 1 mM 0.5 M EDTA pH 8) and centrifuged at 1,316 x g, 4°C for 5 minutes. The cell pellet was resuspended in 400 μL of Lithium Acetate Buffer and used for up to a week.

2.3.3 Transformation of *E. coli* and *S. cerevisiae* competent cells

Transformation of chemically competent *E. coli* cells One Shot^{TM} TOP10 and BL21 – CodonPlus (DE3) – RIL cells was conducted using heat shock. The 50 μL aliquot of the cells were thawed on ice and incubated with 1 – 10 μL of plasmid for 30 minutes. The reaction was heat shocked at 42°C for 45-90 seconds, and then resuspended in 200 μL for of LB media for 1 hour to recover. The entire volume of reaction was spread on LB agar plate with the appropriate antibiotic and incubated at 37°C for 12-16 hours.

The Electrocompetent *E. coli* cells TG1 (Lucigen) and ElectroMAX^{TM} DH5α- E (Invitrogen) were transformed by thawing the cells on ice for 5 minutes before adding 1 μL of DNA to the cells. The reaction was transferred to a 0.1 cm cuvette (Flowgen Bioscience) and Electroporated using a BioRAD GenePulser Xcell^{TM} with the following applied settings: 1.8 kV, 200 Ω, 25 μF. The reaction was resuspended in 1 mL S.O.C medium and transferred to a 15 mL Falcon^{TM} tube (Fisher Scientific). The reaction was incubated at 37°C for 1 hour to recover and 100 μL of the reaction was spread on LB agar and incubated at 37°C for 12-16 hours.

Two co-expression plasmids were transformed into *BCY123 S. cerevisiae* cells to express multiple proteins. Co-expression plasmids pRS-424 and pRS-426 vectors containing the relevant protein encoding sequences were transformed by adding 0.8 μg of each plasmid to 30 μl – 70 μL of competent cells and 20 ng of Salmon Sperm Carrier DNA. The reaction was incubated at 30°C for 15 minutes and then 150 μL of PEGTE (40 % PEG, 10 mM Tris-HCL pH 8 and 1 mM EDTA) was added to the transformation reaction. The reaction was incubated at 30°C for 15 minutes and subsequently incubated at 42°C for 8 minutes. The reaction was incubated on ice for 2 minutes and microcentrifuged using an eppendorf centrifuge-5424 at 18,970 x g for 30 seconds. The pellet was resuspended in 100 μL of sterile water and 50 μL of the cells were plated on selective YEPD agar.

2.3.4 Diagnostic digestion of plasmid DNA

To confirm the protein-encoding gene was present and intact in the plasmid, the DNA sequence was cleaved at the 5' and 3' ends using appropriate restriction enzymes. The reaction, summarised in Table 2.2, was incubated at 37°C for 1 hour. 10 μL of the

Component	Amount
Plasmid	1 μg
Restriction enzyme 1	1 μL
Restriction enzyme 2	2 μL
10x Buffer	2 μL
100x BSA	0.2 μL
H_20	up to 20 μL

Table 2.2 **Restriction digest.** Summarising the contents of the diagnostic digest reaction.

reaction was transferred to a 1% TBE or TAE agarose gel. DNA gel electrophoresis was conducted and the gel was imaged to identify a band corresponding to the size of the protein-encoding gene.

2.3.5 Site-directed mutagenesis

Site directed mutagenesis was conducted using PCR. The mutations were introduced using QuikChange® mutagenesis kit (Agilent) or the Q5® Site-directed mutagenesis kit (NEB). The QuikChange® mutagenesis kit (Agilent), was used to delete the 6x Poly Histidine tag from the N-terminus of Cep3. The PCR reaction was conducted on pRS-426 encoding N-terminally His tagged Cep3 and untagged Skp1, with PfuUltra II Fusion HS DNA Polymerase (Agilent), carried out using a SureCycler 8800 Thermal Cycler (Agilent) with 18 cycles. The reaction and PCR parameters are summarised in Table 2.3 and 2.4. The PCR product was digested with Dpn1 at 37°C for 2 hours to digest the methylated parent plasmid. The resulting product was transformed into chemically competent Top10 *E. coli* cells.

Q5® Site-directed mutagenesis kit was used to insert STREP II tag to the C-terminus of Ctf13 in pRS-424. The PCR was conducted using Q5® Hot Start High-Fidelity DNA Polymerase, on the template plasmid pRS-424 which encodes Sgt1 with C-terminal Strep II tag and Ctf13 C-terminal Calmodulin Binding Peptide (CBP) tag using the settings summarised in Table 2.5 and 2.6. The PCR reaction was treated with Kinase, Ligase and Dpn1 as described by the manufacturer and transformed into Top10 *E. coli* cells. The relevant DNA primers for both experiments are listed in Table 2.7.

Components	**Final Concentration**
Template plasmid	25 ng
Forward mutagenic primer	12.5 ng
Reverse mutagenic primer	12.5 ng
dNTPs	4 μM
Pfu buffer	1x
Pfu	1 unit

Table 2.3 **QuikChange® site directed mutagenesis reaction.** Summary of the contents used for the QuikChange® site directed mutagenesis reaction.

PCR Parameters
Number of cycles 18
Initial denaturation 95°C for 20 sec
Denaturation 95°C for 20 sec
Annealing 55°C for 20 sec
Extension 72°C for 5 min
Final extension 72°C for 3 min

Table 2.4 **QuikChange® site directed mutagenesis PCR parameters.** The table details the PCR settings used for the QuikChange® site directed mutagenesis reaction.

Components	Final Concentration
Template plasmid	10 ng
Forward mutagenic primer	10 μM
Reverse mutagenic primer	10 μM
Q5® Hot start high-fidelity 2x master mix	12.5 μL
Nuclease-free water	up to 25 μL

Table 2.5 **Q5® site directed mutagenesis.** Summary of the contents used for the Q5® site directed mutagenesis reaction.

PCR Parameters
Number of cycles 25
Initial denaturation 98°C for 30 sec
Denaturation 98°C for 10 sec
Annealing 68°C for 20 sec
Extension 72°C for 2.5 min
Final extension 72°C for 2 min

Table 2.6 **Q5® site directed mutagenesis PCR parameters.** Summary of the contents and PCR parameters used for the Q5® site directed mutagenesis reaction.

Primers	Sequence
Cep3 N-terminal 6x Poly-His deletion Forward Primer	5' CACATAAATAAACCATGTTTAAC CGTACCACTCAACTG 3'
Cep3 N-terminal 6x Poly-His deletion Reverse Primer	5' CAGTTGAGTGGTACGGTTAAAC ATGGTTTATTTATGTG 3'
Q5 Ctf13 STREP Forward Primer	5' CACCCGCAGTTCGAAAAATAAA AGCTTAAGTCGACTTTGTTCCC,3'
Q5 Ctf13 STREP Reverse Primer	5' GCTCCAAGCGCTGCCGCCAAGT GCCCCGGAGGATGA 3'

Table 2.7 **Mutagenic primers.** List of mutagenic primers used for both QuikChange® and Q5® site directed mutagenesis.

2.3.6 MultiSAC expression cassette

The MultiSAC expression system, developed by the Nagai group (MRC Laboratory of Molecular Biology, Cambridge), is an *S. cerevisiae* expression system used to co-express multiple proteins for structural and functional studies. The system uses a modified strain of *S. cerevisiae* cells called BCY123 (BCY123 (*MATα pep4*::*HIS3 prb1*::*LEU2 bar1*::*HIS6 lys2*::*GAL1/10GAL4 can1 ade2 trp1 ura3 his3 leu23,112*) where the auxotrophic genes *trp1*, *leu2* and *ura3* have been deleted. The absence of the auxotrophic genes, enables selection for plasmids encoding these deleted gene products (Miller et al., 1998). Two major vacuolar proteases, Proteinase A and Proteinase B, have been knocked out to prevent proteolysis of expressed proteins (Hecht et al., 2014). The cell strain also contains Galactose-inducible *gal4* gene, where large amounts of Gal4 protein is expressed enabling the transcription of multiple genes in high copy number of plasmids (Traven et al., 2006).

The MultiSac system is composed of three different sets of plasmids. It has five pUC based vectors called yKN (Fig 2.1), a second series of vectors called pENT-3C with multiple cloning sites to enable four expression cassettes to be added to the vector and modified pRS-4200 series of vectors (Fig 2.3).

The yKN vector series is composed of five different pUC based vectors called yKNA, yKNB, yKNQ, yKND and yKNC, and these vectors allows for multi-protein expression cassette to be assembled (Fig 2.1). The gene of interest is cloned in between the 5' NcoI restriction site, where the NcoI restriction site encodes the ATG initiation codon and 3' BamHI restriction site, which encodes Glycine – Serine linker. The gene of interest is cloned in between the strong GALGAPDH hybrid promoter whose activity is induced by Galactose and inhibited by Glucose and either a stop codon or TEV protease cleavable CBP tag (yKNC) or hexa-histidine tag (yKND) (Fig 2.1). The yKN plasmid series uses rare restriction enzyme cutter sites such as NotI, RsrII, PacI, AscI and BspQ1. The rare cutter enzymes allow excision of the expression cassette and transfer of the cassette to the pENT-3C Multiple cloning sites (Fig 2.2).

To facilitate the transfer of multiple yeast expression cassettes to the yeast expression vector, a 1.8kb region from pDEST^{TM}-8 was amplified and digested out of the vector. The region consists of ccdB gene, whose gene product is a toxiC, which affect the *E. coli* DNA gyrase, preventing propagation of plasmid in *E. coli*. Next to the ccdB gene is the chloramphenicol resistance gene (Fig. 2.3). These two genes are flanked

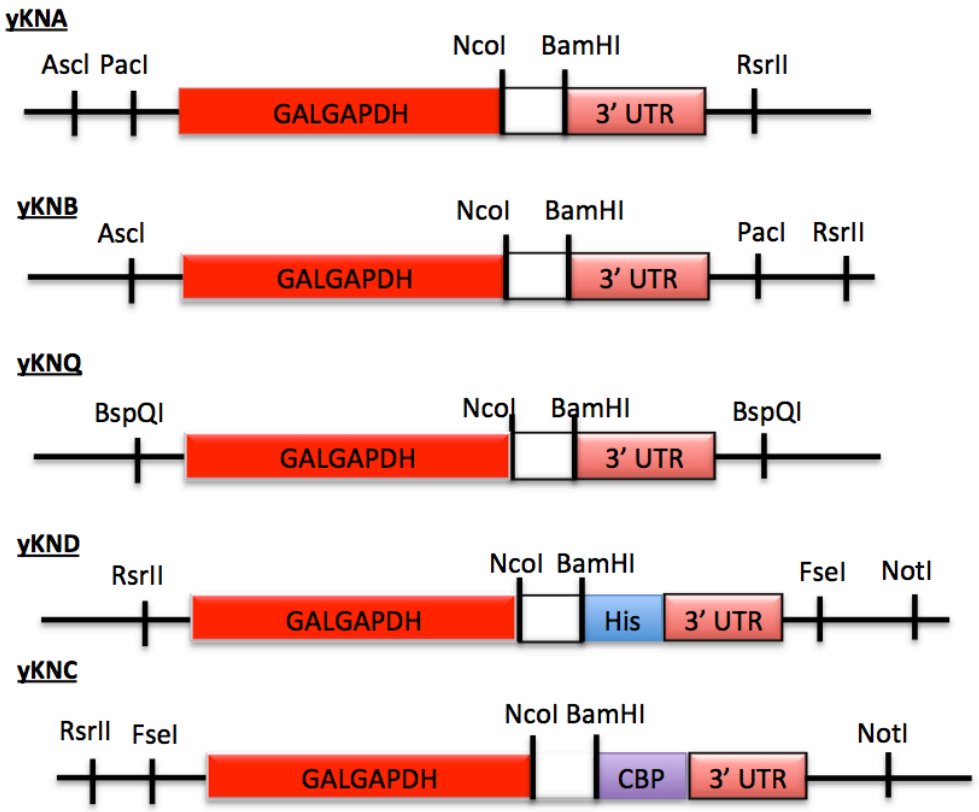


Fig. 2.1 **yKN plasmid series.** Illustration of the yKN plasmids, the restriction sites, the promoter and the insertion site of the protein coding sequence between NcoI and BamHI. The enables the insertion of the protein coding sequence into the plasmid such that it acquires a new promoter and additional affinity purification tags or no tags.

by the attR and attL sites, which facilitate recombination (Fig. 2.3). This region of pDEST^{TM}-8 was digested with XhoI and Bg/II and is cloned into the BamHI and XhoI sites of the pRS-420 series of plasmids. The pRS-420 series in this system consists of three vectors: pRS-424, pRS-425 and pRS-426. The vectors are approximately 7kb in size and it's very difficult to clone different large expression cassettes into the vector directly and instead it's much easier to transfer smaller expression cassettes from the pENT-MSC and yKN plasmid series The pRS-420 series of plasmids are ampicillin resistant and are based on the 2-μ *S. cerevisiae* episomal plasmid (Fig 2.3). The attL and attR site on the pRS-424, pRS-425 and pRS-426 in addition to the attL1 and attL2 sites in pENT-3MSC plasmids enables us to conduct site specific

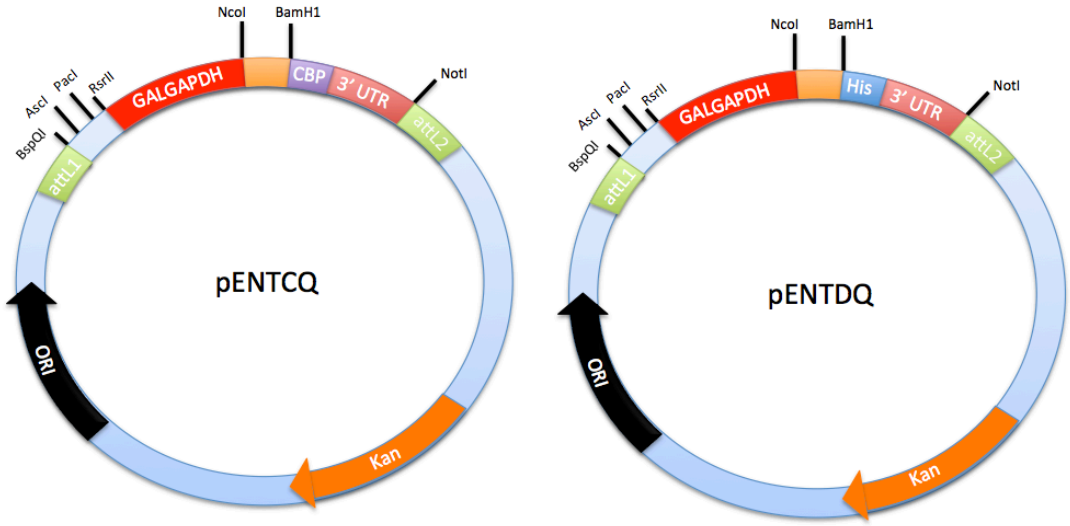


Fig. 2.2 **pENT plasmid series.** Illustration of the pENT plasmids, plasmid features required for plasmid stability and the various insertion sites for multiple protein sequences. The sequence which includes the promoter, gene of interest and possible additional tags, was transferred into the pENT plasmids.

recombination of the expression cassette moving it from pENT-3MSC plasmid to the pRS plasmids in the region of the plasmid which is occupied by ccdB gene and the chloramphenicol resistance gene (Fig. 2.3). The recombination reaction methodology has been developed from bacteriophage λ, and utilises the bacteriophage λ proteins Integrase, excisionase and the *E. coli* encoded integration host factor, to facilitate the exchange using the GATEWAY method (Funk et al., 2002, Katzen, 2007).

As a result, this system enables the user to maintain the three different pRS vectors, due to the availability of BCY123 cell strain with the auxotrophic genes, *TRP1*, *LEU2* and *URA3* knocked out, means we can express up to 12 genes, if we have four proteins encoded per plasmid.

2.3.7 Generating the CBF3-MultiSAC expression cassettes

The MultiSAC expression system was developed and gifted to us by Kiyoshi Nagai (LMB Cambridge). The expression plasmids encoding the proteins were prepared

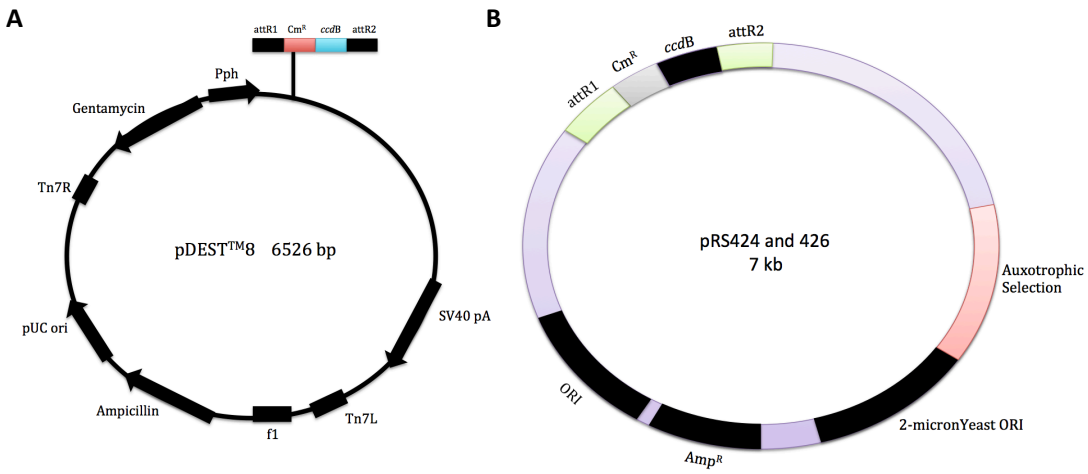


Fig. 2.3 **pDEST^{TM}-8, pRS-424 and pRS-426 plasmids.** A) Illustration of the pDEST plasmid where the region of AttR, chloramphenicol and the toxic ccdB genes are obtained and removed for use in pRS424 and pRS426. (B) Illustration of pRS424 and pRS426, the final expression plasmid generated. The key features the plasmid are listed.

by Dr. Wenjuan Zhang and Jonathan Lucas of the Vaughan lab. To generate the expression cassette, the protein encoding gene was first transferred to a yKN plasmid. Site directed mutagenesis was used to insert NcoI restriction site to the 5' end of the gene and Bam HI restriction site was inserted at the 3' end of the gene. The protein encoding gene was transferred from an *E. coli* plasmid to an empty yKN plasmid by digesting both plasmids for 5 hours at 37°C in a reaction summarised in Table 2.8. Digested vector was purified using 1% Agarose gel electrophoresis (Fisher Miller Reagents) and the Gel extraction kit (QuiagenTM). The digested yKN vector was diluted to a concentration that is smaller than the insert using water and the following three-ligation reactions were conducted at room temperature for 30 minutes (Table 2.9).

The ligations reactions were transformed and the plasmids were purified and sequenced. To transfer the protein expression cassette to pENT vector, both the yKN vector containing the protein coding sequence and the entry vector pENT were digested with the restriction enzyme BspQ1 at 50°C for 3 hours. Both digestions were purified using 1% Agarose gel electrophoresis (Fisher Miller Reagents) and the Gel extraction kit (Quiagen^{TM}) using the protocol and reagents provided. The yKN digested product and

Component	Amount
Plasmid	1 μg
BspQ1	1 μL
NEB buffer 3.1	1 μL
Total Volume	10 μL

Table 2.8 **Gene transfer to yKN plasmids.** Reaction volume and the components used to transfer protein-encoding gene from an *E. coli* plasmid to a yKN plasmid.

	Control	Reaction 1	Reaction 2	Reaction 3
Protein coding sequence	0.0 μg	0.5 μg	1.0 μg	2.0 μg
yKN vector	0.5 μg	0.5 μg	0.5 μg	0.5 μg
10x T4 DNA ligase buffer	1x	1x	1x	1x
T4 DNA ligase	1 unit	1 unit	1 unit	1 unit
Total Volume	10 μL	10 μL	10 μL	10 μL

Table 2.9 **Summary of ligation of gene to yKN vector.** This table summarises the components and the subsequent volumes of each component used to conduct the ligation of protein coding gene into yKN vector.

pENT digested product were diluted, such that the pENT was at a lower concentration than yKN, using water and ligated using the following reaction were conducted at room temperature for 30 minutes (Table 2.10).

The ligation reactions were transformed and the plasmids were purified. The pENT-CQ protein coding sequences were then transferred to the destination vectors pRS-424 or pRS-426 plasmids using the GATEWAY reaction summarised in Table 2.11. The reaction was incubated at 25°C for 16 hours and the reaction was terminated by the

	Control	Reaction 1	Reaction 2	Reaction 3
yKN digestion product	0.0 μg	0.5 μg	1.0 μg	2.0 μg
pENT digestion product	0.5 μg	0.5 μg	0.5 μg	0.5 μg
10x T4 DNA ligase buffer	1x	1x	1x	1x
T4 DNA ligase	1 unit	1 unit	1 unit	1 unit
Total Volume	10 μL	10 μL	10 μL	10 μL

Table 2.10 **Summary of ligation of yKN digestion product to a pENT vector.** This table summarises the components and the subsequent volumes of each component used to conduct the ligation of protein coding gene from yKN vector to the pENT vector.

Component	**Amount**
pENT-CQ plasmid	100 ng
pRS-424/pRS-426 plasmid	200 ng
Buffer (10 mM Tris pH 8, 0.2 mM EDTA pH 8)	5 μL
Invitrogen Gateway LR Clonase Enzyme mix	2 μL
Total Volume	10 μL

Table 2.11 **Summary of the components and the quantities for the gateway reaction.** This table summarises the components and the subsequent volumes or amounts of each component used to conduct the gateway reaction to transfer the protein coding sequence and additional plasmid elements to the pRS-424 or pRS-426 plasmids. .

addition of 1 μL Proteinase K solution (Invitrogen) by incubating at 37°C for 10 minutes. The product was transformed using electroporation.

2.4 Protein expression in *E. coli*

The proteins calmodulin and Cep3 were both expressed in *E. coli* using previously established methodologies. Calmodulin was expressed in order to purify and make calmodulin bound resin for affinity chromatography. Cep3 was expressed in *E. coli* because two labs had previously established a robust protocol to express and purify high yields of Cep3 (Bellizzi et al., 2007, Purvis and Singleton, 2008). Cep3 was expressed in order to purify large quantities of protein for biochemical assays.

2.4.1 Expression of Calmodulin

Calmodulin encoded in the vector pET-15b, gifted by Professor Kiyoshi Nagai (LMB), was transformed into BL21– CodonPlus (DE3) cells (Agilent). A colony from the LB agar plate with ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL) was streaked on to 6 LB agar plates with the same antibiotics, which was incubated at 37°C for 12-16 hours. The colonies were resuspended in 6x 100 mL flasks with 10 mL of 2x TY media. The flasks were incubated at 37°C for 12-16 hours. 5 mL from each flask was used to inoculate 12 flasks if 1 litre of 2x TY media. The cell cultures were grown at 37°C until $OD_{600_{nm}}$ of 0.8 – 1.0 is reached. The cells were induced with 0.4 mM IPTG. The cell cultures were grown for 3-5 hours. The cells were harvested using J6-MI Centrifuge

(Beckman Coulter) at 4540 x g at 4°C for 20 minutes and resuspended in lysis buffer (25 mM Tris-Cl pH 7.4, 1 M NaCl, 10 mM imidazole pH 8, 10 mM β-mercaptoethanol). Protein was expressed across 12 litres of cell culture.

2.4.2 Expression of Cep3

Cep3 protein was encoded in the pET-28a vector. The plasmids were transformed into BL21– CodonPlus (DE3) cells (Agilent). A colony from the LB agar plate with kanamycin (50 μg/mL) and chloramphenicol (34 μg/mL) for each protein was used to inoculate separate 150 mL of TB media. The cell cultures were grown at 37°C for 12-16 hours. 6 flasks with 1 litre of TB media were inoculated with 10 mL of the grown cell culture for each appropriate plasmid and the flasks were incubated at 37°C until $OD_{600_{nm}}$ of 0.8 was reached. The cells were induced with 0.5 mM IPTG. The flasks expression of full length Cep3 had 100 μM of $ZnCl_2$ supplemented. The induced cell cultures were incubated at 18°C for 12 hours. The cells were harvested using J6-MI Centrifuge (Beckman Coulter) at 4540 x g at 4°C for 15 minutes and resuspended in lysis buffer (60 mM Tris pH 8, 0.5 M NaCl, 0.5% Tween 20, 8% (w/v) glycerol, 5 mM β-mercaptoethanol and 1 protease inhibitor tablet (Roche) per 40 mL of buffer. Protein was expressed across 6 litres of cell culture.

2.5 Protein expression in *S. cerevisiae*

Previous attempts in the lab to establish a methodology to express the CBF3 complex in *E. coli* were unsuccessful. This is because the CBF3 complex assembly is dependent on Sgt1 and Hsp90. In order to express the CBF3 complex, it was determined that Sgt1 had to be co-expressed with the subunits. Hsp90 was not co-expressed, but it's speculated that the cell may have enough endogenously expressed Hsp90 to facilitate complex formation. Hence why in order to purify the complex for further study, protein expression of the CBF3 complex was done in *S. cerevisiae* cells.

2.5.1 Protein expression of the CBF3 core complex

Truncated CBF3 core complex

The truncation was carried out on Cep3 were residues 1-46 were removed from the Cep3 N-terminus. The pRS-424 plasmid encoding Ctf13 and Sgt1 were co-transformed with pRS-426 plasmid encoding Cep3 and Skp1 into BCY123, where the colonies from the transformation were then resuspended in 300 mL YEPD media (6.9 mg/mL Yeast nitrogen base without amino acids (Formedium), 11 mg/mL Casamino Acids (Formedium), 55 μg/mL Adenine Sulphate, (Formedium) 55 μg/mL L- tyrosine (Sigma)) supplemented with 2 % D (+) – Raffinose (Formedium) and 100 μg/mL Ampicillin. The starter culture was grown at 30°C for 24 hours. The starter culture was added to 800 mL YEPD media, supplemented with 1 % D (+) – Raffinose (Formedium) and 100 μg/mL Ampicillin, $OD_{600_{nm}}$) of 0.25. The cultures were grown at 30°C until $OD_{600_{nm}}$) of 0.8 – 1.0 was reached. The cells were induced with 2 % D (+) – Galactose. The cells were incubated for 16 – 18 hours at 30°C. The cells were harvested by centrifugation using a Beckman Coulter J6-MI at 4540 x g at 4°C for 30 minutes. The cell pellet was resuspended with MiliQ filtered water and centrifuged at 4540 x g at 4°C for 30 minutes. The cells were resuspended in 2x Lysis Buffer (25 % Glycerol, 50 mM Tris-Cl pH 8.0, 500 mM NaCl, 2 mM Magneisium acetate, 2 mM Imidazole, 4 mM $CaCl_2$, 0.2% IGEPAL CA-630 (Sigma), 40 mM Sodium fluroide, 4 μg/mL DNaseI, 4 mM DTT and 4 Protease Inhibitor Cocktail tablets (Roche)). The cells were added to liquid nitrogen drop by drop to make 'droplets' and stored at – 80°C. For each expression, 2 starter cultures were prepared with 6 selective YEPD plates containing transformants. The two starter cultures were used to inoculate 24 flasks of 800 mL YEPD with additives.

Full length CBF3 core complex

The protocol for expression of the full-length complex was the same as the truncated CBF3 core complex with two changes. First the cell culture was supplemented with 100 μM $ZnCl_2$ prior to induction of protein expression. After harvesting the cells, they were also resuspended into a different buffer. The cells were resuspended in a 2x lysis buffer composed of 25 % Glycerol, 40 mM Tris-Cl pH 8.0, 500 mM NaCl, 2 mM Magneisium acetate, 0.2 % IGEPAL CA-630 (Sigma), 100 mM Sodium fluoride, 4 μg/mL DNase I, 5 mM β-mercaptoethanol and 4 Protease Inhibitor Cocktail tablets, and then added to liquid nitrogen drop by drop to make 'droplets' which are stored at – 80°C.

2.5.2 Protein expression of Ndc10

To express Ndc10 alone in *S. cerevisiae*, the cells were transformed with a pRS-424 plasmid encoding Ndc10 and an empty pRS-426 plasmid. This was done to ensure the cells are provided with both uracil and tryptophan, which have been knocked out to act as a selection marker, to meet the cells auxotrophic needs. The transformed colonies were resuspended in 300 mL YEPD media (6.9 mg/mL Yeast nitrogen base without amino acids (Formedium), 11 mg/mL Casamino Acids (Formedium), 55 μg/mL Adenine Sulphate, (Formedium) 55 μg/mL L- tyrosine (Sigma)) supplemented with 2 % D (+) – Raffinose (Formedium), 100 μg/mL Ampicillin and 0.1% Glucose. The starter culture was grown at 30°C for 24 hours. The starter culture was added to 800 mL YEPD media, supplemented with 1 % D (+) – Raffinose (Formedium), 100 μg/mL Ampicillin and 0.1% Glucose (Formedium), to $OD_{600_{nm}}$of 0.25. The cultures were grown at 30°C until $OD_{600_{nm}}$ of 0.8 – 1.0 was reached. The cells were induced with 2 % D (+) – Galactose. The cells were incubated for 16 – 18 hours at 30°C. The cells were harvested by centrifugation at 4540 x g at 4°C for 30 minutes. The cell pellet was resuspended with MiliQ filtered water and centrifuged at 4540 x g at 4°C for 30 minutes. The cells pellet was resuspended in 2x Lysis Buffer (25 % Glycerol, 40 mM Tris-Cl pH 8.0, 500 mM NaCl, 2 mM Magnesium acetate, 1 mM EDTA pH 8, 0.2 % IGEPAL CA-630 (Sigma), 40 mM Sodium fluoride, 4 μg/mL DNaseI, 3 mM DTT and 4 cOmpleteTM, Mini, EDTA-free Protease Inhibitor Cocktail tablets (Roche)). The cells were added to liquid nitrogen drop by drop to make 'droplets' and stored at – 80°C. For each expression, 2 starter cultures were prepared with 6 selective YEPD plates containing transformants. The two starter cultures were used to inoculate 24 flasks of 800 mL YEPD with additives.

2.5.3 Protein expression of truncated CBF3 complex

The CBF3 complex expressed had truncations at the N-terminus of Cep3 from residues 1-46 for stability and Ndc10 was truncated at the C-terminus from residue 702 – 956. The pRS-424 plasmid encoding Ctf13 and Sgt1 were co-transformed with pRS-426 plasmid encoding Cep3, Ndc10 and Skp1 into BCY123, where the colonies from the transformation were then resuspended in 300 mL YEPD media (6.9 mg/mL Yeast nitrogen base without amino acids (Formedium), 11 mg/mL Casamino Acids (Formedium), 55 μg/mL Adenine Sulphate, (Formedium) 55 μg/mL L- tyrosine (Sigma))

supplemented with 2 % D (+) – Raffinose (Formedium), 0.1% glucose (Formedium) and 100 μg/mL Ampicillin (Formedium). The starter culture was grown at 30°C for 24 hours. The starter culture was added to 800 mL YEPD media, supplemented with 1 % D (+) – Raffinose (Formedium) and 100 μg/mL Ampicillin, to an $OD_{600_{nm}}$nm) of 0.25. The cultures were grown at 30°C until $OD_{600_{nm}}$ of 0.8 – 1.0 was reached. The cells were induced with 2 % D (+) – Galactose. The cells were incubated for 16 – 18 hours at 30°C. The cells were harvested by centrifugation using a Beckman Coulter J6-MI at 4540 x g at 4°C for 30 minutes. The cell pellet was resuspended with MiliQ filtered water and centrifuged at 4540 x g at 4°C for 30 minutes. The cells were resuspended in 2x Lysis Buffer (25 % Glycerol, 50 mM Tris-Cl pH 8.0, 500 mM NaCl, 2 mM Magnesium acetate, 2 mM Imidazole, 4 mM $CaCl_2$, 0.2 % IGEPAL CA-630 (Sigma), 40 mM Sodium fluoride, 4 μg/mL DNaseI, 4 mM DTT and 4 Protease inhibitor cocktail tablets (Roche)). The cells were added to liquid nitrogen drop by drop to make 'droplets' and stored at – 80°C. For each expression, 2 starter cultures were prepared with 6 selective YEPD plates containing transformants. The two starter cultures were used to inoculate 24 flasks of 800 mL YEPD with additives.

2.6 Protein purification

2.6.1 *E. coli* cell lysis

The *E. coli* cells were lysed using an Avestin Emulsiflex- C5 (Wolfs Lab). The cell pellet was resuspended in the appropriate lysis buffer and then applied to the Emulsiflex. The cells were lysed using 3 cycles at a pressure of 15,000 Psi. For purification of Calmodulin, the lysate was cleared by centrifugation at 38,465 x g for 1 hour at 4°C using a Sorvall RC 6+ (Thermo Scientific) centrifuge. For Cep3, the lysate was cleared at 167,424 x g for 45 minutes at 4°C using a Optima^{TM} L-90K Ultracentrifuge (Beckman Coulter).

2.6.2 *S. cerevisiae* cell lysis

The cell lysis was conducted using a 6870 Freezer/Mill® Cryogenic Grinder. The cell 'droplets' were placed inside a sealed vial, up to a third of the vial volume, with a steel

Plasmid	**Proteins encoded**	**Vector and selection marker**	**Protein tag**
Cep3ΔN-Skp1	Cep3 (Residues 47 – 608) Skp1	pRS-426 *TRP*	Cep3 (N-term 6x His)
Cep3 - Skp1	Cep3 Skp1	pRS-426 *TRP*	Cep3 (N-term 6x His tag)
Cep3 - Skp1	Cep3 Skp1	pRS-426 *TRP*	No tags
Sgt1 - Ctf13	Sgt1 Ctf13	pRS-424 *URA*	Sgt1 (C-term Strep II) Ctf13 (C-term CBP)
Sgt1 -Ctf13	Sgt1 Ctf13	pRS-424 *URA*	Sgt1 (C-term Strep II) Ctf13 (N-term 6x His)
Sgt1 - Ctf13	Sgt1 Ctf13	pRS-424 *URA*	Sgt1 (C-term Strep II Ctf13 (C-term 6x His)
Sgt1 - Ctf13	Sgt1 Ctf13	pRS-424 *URA*	Sgt1 (C-term Strep II) Ctf13 (C-term CBP-Strep II)
Cep3ΔN - Skp1 - Ndc10 D1-2	Cep3 (Resisudes 47 - 608) Skp1 Ndc10 (Residues 1 -554)	pRS-426 *TRP*	Cep3 (N-term 6x His) Ndc10 (C-term Strep II)
Cep3ΔN - Skp1 - Ndc10 D1-3	Cep3 (Residues 47 - 608) Skp1 Ndc10 (Residues 1 -701)	pRS-426 *TRP*	Cep3 (N-term 6x His) Ndc10 (C-term Strep II)
Ndc10 D1-2	Ndc10 (Residues 1 - 554)	pRS-424 *URA*	Ndc10 (C-term Strep II)
Ndc10 D1-3	Ndc10 (Residues 1 - 701)	pRS-424 *URA*	Ndc10 (C-term Strep II)

Table 2.12 **List of Protein expression plasmids used.** This table summarises the all of the different plasmids which were used to express the complex and the subunits in *S. cerevisiae*.

impactor. This was placed within the grinder which is operated with liquid nitrogen. The cells were ground for four cycles at a rate of 15 cps. The initial run had a pre-cool time of 5 minutes and subsequent runs had a pre-cool time of 1 minute. The run time for each cycle is 2 minutes, with a 2 minute pause between each cycle. The pulverised cells form a powder, which is then thawed and resuspended in appropriate 2x lysis buffer. The pH of the lysate is adjusted to 8 using 1M Tris-Cl pH 11 and the lysate is cleared via ultracentrifugation using a Optima^{TM} L-90K Ultracentrifuge (Beckman Coulter) at 167,424 x g at 4°C for 90 minutes.

2.6.3 Ammonium Sulphate precipitation

This was used to purify Cep3 only. The ammonium sulphate (Sigma) was gradually added to the cleared supernatant at 0°C to achieve 65% ammonium sulphate. Upon addition of ammonium sulphate, the mixture was left to stir slowly for 2 hours. The precipitated protein was recovered by ultracentrifugation using a Optima^{TM} L-90K Ultracentrifuge (Beckman Coulter) at 167,424 x g for 45 minutes at 4°C. The supernatant was discarded and the buffer from the next purification step was added to pellet with a volume which is equivalent to the supernatant, to resuspended the cell pellet.

2.6.4 Calmodulin affinity purification

This is an affinity based chromatography step where Calmodulin attached to the resin interacts with a short calmodulin binding peptide (CBP) derived from the C-terminal fragment from muscle myosin light chain kinase. When the calcium concentration is low, the peptide displays strong affinity for Calmodulin. Stripping off the calcium from the calmodulin induces conformational change in Calmodulin, which results in the release of the protein (Torok and Trentham, 1994, Waxham et al., 1998). The calmodulin affinity resin was prepared in the lab and described in Section 2.6.9

A small volume of 900 μL of Calmodulin resin was incubated with the cleared lysate for 12 hours at 4°C. The pre-equilibrated resin was washed with 'CAM wash buffer' composed of 10 mM Tris-Cl pH 8, 500 mM NaCl, 1 mM Magnesium acetate, 1 mM Imidazole, 2 mM Calcium chloride, 10 mM β-mercaptoethanol. The protein was eluted

in 'CAM elution buffer' composed of 10 mM Tris-Cl pH 8, 500 mM NaCl, 1 mM Magnesium acetate, 1 mM Imidazole, 10 mM β-mercaptoethanol and 4 mM EGTA.

2.6.5 Immobilised Metal Affinity Chromatography

Immobilised Metal Affinity Chromatography (IMAC) was used to interact and bind proteins tagged with 6 histidines residues. This purification as done using both pre-packed columns or with resin. In terms of resin, both TALON® cobalt resin (Clontech) and Ni SepharoseTM High performance resin (GE Healthcare) were used to try purify the full length CBF3 core complex. The pre-equilibrated resin was washed with 20 resin volume of wash buffer containing 20 mM Imidazole. The protein was eluted from the resin using 10 resin volumes of elution buffer containing 500 mM Imidazole.

Alternatively, metal affinity purification was also conducted using 1 or 5 mL Nickel HisTRAP^{TM} HP (GE Healthcare). The protein was applied to the pre-equilibrated column using a Peristaltic Pump P-1 (GE Healthcare) or a 50 mL Superloop (GE Healthcare). The columns were next transferred to a ÄKTApurifier (GE Healthcare) or ÄKTAprime plus (GE Healthcare) where the proteins were eluted across a gradient reach maximum concentration of 500 mM Imidazole.

2.6.6 Ion exchange chromatography

Anion exchange chromatography was used to prepare a number of proteins and protein complexes. This purification was conducted using a 1 mL Resource^{TM} Q (GE Healthcare) pre-equilibrated on an ÄKTApurifier (GE Healthcare) with 5 column volumes of buffer containing no NaCl, then washed with 10 column volumes of 1 M NaCl followed by a second 5 column volume wash with buffer containing 150 or 200 mM NaCl depending on the sample. The protein samples were applied to the column using a 50 mL Superloop (GE Healthcare). The proteins were eluted across a specified linear gradient or step elution.

2.6.7 Heparin chromatography

The Heparin attached to the resin has strong affinity for DNA binding proteins and therefore the Heparin chromatography step can be used to separate DNA binding proteins from non-DNA binding proteins. This step was conducted with a 1 mL HiTrap® Heparin HP (GE Healthcare). The column was pre-equilibrated on an ÄKTApurifier (GE Healthcare) with 5 column volumes of buffer containing no NaCl, then washed with 10 column volumes of buffer containing 1 M NaCl followed by a second 5 column volume wash with buffer containing 150 or 200 mM NaCl depending on the sample. The protein samples were applied to the column using either a 50 mL Superloop (GE Healthcare) or Peristaltic Pump P-1 (GE Healthcare). The sample was eluted across either a linear gradient or step elution, depending on the sample.

2.6.8 Gel filtration

For samples were there is a large concentration of proteins such as Cep3, or when Gel filtration is required in the middle of the work flow rather than the final step, the columns HiLoad 16/60 Superdex 200 PG (GE Healthcare) and HiLoad 16/60 Superdex 75 PG (GE Healthcare) were used according to their specifications, injecting 1-5 mL of protein on to the column on an ÄKTAprime plus (GE Healthcare). These columns were equilibrated with 1.5 column volumes of size exclusion buffer, and the proteins were eluted across 1 column volume using the same buffer.

For samples where obtaining high yields was not feasible or samples prepared for structural studies, the Superdex increase 200 10/300 GL (GE Healthcare) was used on an ÄKTApurifier (GE Healthcare). A small volume of 100 – 500 μL was applied on to the column. The column was equilibrated with 1.5 column volumes of the size exclusion buffer and the proteins were eluted across 1 column volume using the same buffer.

2.6.9 Purification of Calmodulin and Calmodulin resin preparation

E. coli cells containing expressed Calmodulin were sonicated for a total of 5 minutes, with the pulse on for 2 seconds and off for 4 seconds at 40% amplitude for total of 3 cycles. After sonication, the lysate was pumped through an Avestin Emulsiflex – C5 (Wolfslab) at 15,000 Psi once. The lysate was cleared by ultracentrifugation at 142,032 x g for 30 minutes. The cleared lysate was loaded onto two 5 mL Nickel HisTrap^{TM} (GE Healthcare), pre-equilibrated with Buffer A (25 mM Tris-Cl pH 7.4, 500 mM NaCl, 10 mM Imidazole, 10 mM β-mercaptoethanol) using a Peristaltic Pump P-1 (GE Healthcare). The Calmodulin was eluted from the column using a linear gradient from Buffer A to Buffer B (25 mM Tris-Cl pH 7.4, 500 mM NaCl, 250 mM Imidazole, 10 mM β-mercaptoethanol). The protein fractions are then dialyzed for 12 hours into Q Buffer A (25 mM Tris-Cl pH 7.4, 25 mM NaCl, 5 mM $CaCl_2$, 5 mM β-mercaptoethanol). The dialyzed sample was loaded on to 20 mL HiLoad^{TM} 16/10 Q-Sepharose^{TM} High Performance using a 50 mL Superloop (GE Healthcare). The protein was eluted using a linear gradient from Q Buffer A to Q Buffer B (25 mM Tris-Cl pH 7.4, 1M NaCl, 5 mM $CaCl_2$, 5 mM β-mercaptoethanol). The protein fractions were concentrated using an Amicon® Ultra 15 mL Centrifugal Filters (Merck) and applied to a HiLoad 16/6-Superdex 200 PG (GE Healthcare). The column was equilibrated and the protein was eluted in the buffer containing 500 mM NaCl, 0.1 M $NaHCO_3$ pH 8.3. The coupling of the calmodulin to cyanogenbromide-activated Sepharose was done as instructed by the manufacture manual (GE healthcare).

2.6.10 Purification of Cep3

The cell pellet was pumped through Avestin Emulsiflex – C5 (Wolfslab) at 15,000 Psi 3 times. The lysate was cleared by centrifugation at 78,378 x g for 45 minutes at 4°C. Ammonium Sulphate precipitation was conducted by adding Ammonium Sulphate dry powder to the lysate to achieve 65% Ammonium Sulphate at 0°C. After stirring the mixture for 2 hours, the sample was centrifuged at 167424 x g for 45 minutes at 4°C. The pellet was resuspended in Buffer A (20 mM Tris- Cl pH 8, 500 mM NaCl, 20 mM Imidazole) using the same volume as the supernatant that was removed. The resuspension was loaded on to a 5 mL Nickel HisTrap^{TM} (GE Healthcare), pre-equilibrated with Buffer A using a Peristaltic Pump P-1 (GE Healthcare). The protein

was eluted using a linear gradient from Buffer A to Buffer B (20 mM Tris- Cl pH 8, 500 mM NaCl, 500 mM Imidazole, 2 mM DTT). The N-terminal His tag was cleaved at 4°C for 12 hours and then 3 hours at room temperature using thrombin (GE Healthcare). The sample was then loaded on to 1 mL Nickel HisTrap^{TM} (GE Healthcare), pre-equilibrated with Buffer A. The flow through was then concentrated using an Amicon® Ultra 15 mL Centrifugal Filters (Merck) and applied to a HiLoad 16/6- Superdex 75 PG (GE Healthcare). The column was equilibrated and the protein as eluted in Buffer containing 20 mM Tris-Cl pH 8, 200 mM NaCl and 2 mM DTT. When purifying the full length Cep3, all buffers also contain 10 μM of $ZnCl_2$.

2.6.11 Purification of CBF3ΔN core complex

The protocol established By Dr. Wenjuan Zhang in the lab involved co-expression of two constructs (Table 2.12) which were expressed in 24 litres of *S. cerevisiae* cell culture and harvested (Section 2.5.1), and re-suspended in 2x lysis buffer (25 % Glycerol, 50 mM Tris-Cl pH 8.0, 500 mM NaCl, 2 mM Magnesium acetate, 2 mM Imidazole, 4 mM $CaCl_2$, 0.2 % IGEPAL CA-630 (Sigma), 40 mM NaF, 4 μg/mL DNase I, 4 mM DTT and 6 Protease Inhibitor Cocktail tablets (Roche), before being stored as droplets in liquid nitrogen. The cell were lysed and the lysate was cleared (Section 2.6.2). The first step of the purification involved incubation of the cleared lysate with calmodulin resin (2.6.4). The protein eluted was concentrated to 8 mL using an Amicon® Ultra Centrifugal filter (Merck) and dialysed at 4°C for 12-16 hours into the buffer: 20 mM Tris-Cl pH 8, 500 mM NaCl, 5 mM Imidazole, 10 % Glycerol, 10 mM β-mercaptoethanol. The protein was loaded on to a pre-equilibrated 1 mL Nickel HisTrap^{TM} HP (GE Healthcare). The column was washed with Nickel A buffer (20 mM Tris-Cl pH 8, 500 mM NaCl, 20 mM Imidazole, 10 mM β-mercaptoethanol). The protein was eluted with a linear gradient going from 10-25 % of imidazole in the buffer: 20 mM Tris-Cl pH 8, 500 mM NaCl, 500 mM Imidazole, 10 mM β-mercaptoethanol.

Plasmid	Proteins	Vector and Selection marker	Protein tag
1	Cep3 (Residue 47-608) Skp1	pRS-426 TRP	Cep3 (N-term 6x His)
2	Sgt1 Ctf13	pRS-424 URA	Sgt1 (C-term Strep II) Ctf13 (C-term CBP)

Table 2.13 **CBF3ΔN core complex plasmids.** Plasmids constructs used to purify the CBF3δN core complex.

The pooled protein was concentrated using an Amicon® Ultra Centrifugal filter (Merck) and the sample was diluted using the QA buffer (20 mM Tris-Cl pH 8, 100 mM NaCl, 5 mM EDTA pH 8, 10 % Glycerol, 2 mM DTT) to dilute the salt concentration in the sample to 150 mM. The diluted sample was loaded onto 1 mL Resource^{TM}Q. The column was washed with QA buffer and the protein was eluted in buffer QB (20 mM Tris-Cl pH 8, 1 M NaCl, 5 mM EDTA pH 8, 10 % Glycerol, 2 mM DTT) with a linear gradient of 5.6 %-22 % of buffer QB. The Superdex 200 10/300 GL (GE Healthcare) was used in the final step of the purification using method described in 4.14 in buffer containing: 15 mM Tris-Cl pH 8, 200 mM NaCl, 2 mM DTT. The SDS-PAGE and chromatogram shows three bands for the truncated Cep3, Ctf13 and Skp1 (Fig 2.4). After gel filtration, the purification yielded a complex with three subunits, with no visible contaminants (Fig 2.4.D).

2.7 SDS-PAGE

The sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was conducted using pre-cast Bolt^{TM} 4-12% Bis-Tris Plus Gels (Invitrogen) inside Mini Gel Tank (Invitrogen). The sample was electrophoresed at 200 V for 30 minutes. The protein bands were visualised by staining the gel with InstantBlueTM Protein Stain (Expedeon) or Coomasie Blue.

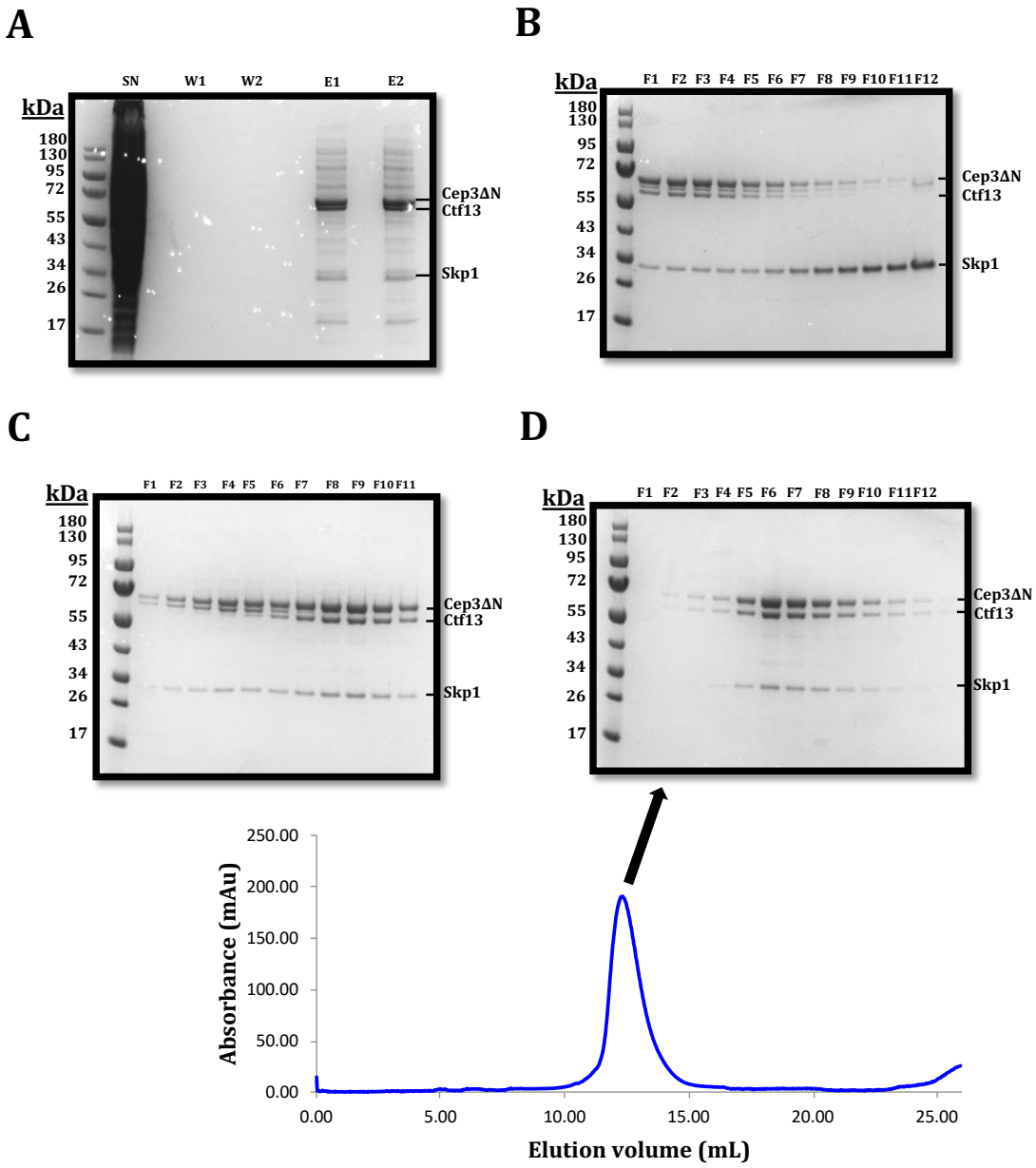


Fig. 2.4 **Purification of CBF3ΔN core complex.** (A) SDS-PAGE after incubation with calmodulin resin. SN represents the supernatant, W1 is wash 1 and E1 is elution1. The Δ represents residues 1-46, which were truncated from Cep3. (B) SDS-page after nickel affinity chromatography. The F1 is fraction 1. (C) SDS page after the anion exchange chromatography step. (D) SDS-PAGE and chromatogram after gel filtration.

2.8 Electrophoretic Mobility Shift Assay

The Electrophoretic Mobility Shift Assay (EMSA) assay was conducted on the full-length CBF3 core complex, truncated CBF3 core complex (CBF3ΔN), and the DNA binding region of Ndc10 (Ndc10 domains 1-2). Each assay was conducted with the 10x reaction buffer (Table 2.14), using either 5' FAM labelled 33 bp CDEIII or 56 bp CDEIII oligonucleotides (Table 2.15).

10x Reaction Buffer
250 mM Hepes pH 8
2 M KCl
20 mM DTT
100 mM MgCl2
0.2% NP-40

Table 2.14 **10x EMSA reaction buffer.** Components of the 10x reaction buffer used to conduct the incubation of protein and DNA.

Name	**Sequence**
33 bp CDEIII Forward sequence	5' [6FAM]AAATATTAGTGTATTTGATTTCCGAAAGTTAAA 3'
33 bp CDEIII Reverse sequence	5' [6FAM]TTTAACTTTCGGAAATCAAATACACTAATATTT 3'
56 bp CDEIII Forward sequence	5' [6FAM]TATTAGTGTATTTGATTTCCGAAAGTTAAAAAAGAAATAGTAAGAAATATATATTT 3'
56 bp CDEIII Reverse sequence	5' [6FAM]AAATATATATTTCTTACTATTTCTTTTTTAACTTTCGGAAATCAAATACACTAATA 3'

Table 2.15 **List of DNA sequences used for EMSA reaction.** The sequence of the individual oligonucleotides which were annealed together to form the DNA duplex, used to study protein or protein complex binding.

CBF3ΔN core complex

The EMSA assay for the CBF3ΔN core complex with fluorescently labelled 33 bp CDEIII and 56 bp CDEIII was assembled as described in the tables (Table 2.16). The assembled reactions were incubated at room temperature for 40 minutes. The 10% glycerol in the buffer was present to aid sample loading on the gels. The EMSA assay was conducted with 3-12% Bis-Tris Novex polyacrylamide gel (Invitrogen) using 1x Native PAGE running buffer (Invitrogen). The gel electrophoresis was done using a

X-Cell SureLock^{TM} Mini-Cell Electrophoresis system. Prior to sample loading, the gel was electrophoresed with 1x Native PAGE running buffer (Invitrogen), on ice, at 4 mA for 60 minutes. The 20 μL reaction was loaded on to the gel. The sample was electrophoresed on ice, at 150 V for 108 minutes, covered with aluminium foil, to prevent photobleaching of the DNA fluorescent label.The sample was imaged using the FLA-3000 imaging system (Fujifilm), at the excitation wavelength of 472 nm. The band shift was determined by comparing the position of the shifted DNA band with the position of the DNA band in the DNA only control.

Component	**DNA : CBF3ΔN Core Complex**				
	1:0	**0:4**	**1:1**	**1:2**	**1:4**
DNA	3.2 μM	–	3.2 μM	3.2 μM	3.2 μM
CBF3ΔN Core Complex	–	12.8 μM	3.2 μM	6.4 μM	12.8 μM
10x Reaction Buffer	1x	1x	1x	1x	1x
Glycerol	10%	10%	10%	10%	10%
Total Volume	20 μL	20 μL	20 μL	20 μL	20 μL

Table 2.16 **EMSA reaction of the CBF3ΔN core complex with 33 bp and 56 bp CDEIII.** Components and the volumes used to make a 20 μL reaction for the EMSA assay of CBF3ΔN with either 33 bp or 56 bp CDEIII.

The CBF3 core complex

The EMSA assay for the CBF3 core complex with fluorescently labelled 33 bp CDEIII and 56 bp CDEIII was assembled without λ protein phosphatase treatment (Table 2.17) and with λ protein phosphatase treatment (Table 2.18). The assembled reaction, similar to the EMSA reactions of the CBF3ΔN core complex (Table 2.16), with the addition of 10 μM $ZnCl_2$ to stabilise the Zn_2Cys_6 cluster domain of Cep3. The EMSA assay conducted with *Lambda* protein phosphatase (Table 2.18) contained 200 units of phosphatase as well as 1x Protein MetalloPhosphatases (PMP) buffer (NEB) and 1 mM $MnCl_2$, to aid the dephosphorylation reaction. The assembled reactions were incubated at 30°C for 40 minutes. The EMSA assay was then conducted as described for the the CBF3ΔN core complex EMSA assay. The sample was imaged and the band shift was determined as described for the CBF3ΔN core complex EMSA assay.

Component	**DNA : CBF3 Core Complex**				
	1:0	**0:4**	**1:1**	**1:2**	**1:4**
DNA	1.25 μM	–	1.25 μM	1.25 μM	1.25 μM
CBF3 Core Complex	–	5.00 μM	1.25 μM	2.50 μM	5.00 μM
10x Reaction Buffer	1x	1x	1x	1x	1x
$ZnCl_2$	10 μM	10 μM	10 μM	10 μM	10 μM
Glycerol	10%	10%	10%	10%	10%
Total Volume	20 μL	20 μL	20 μL	20 μL	20 μL

Table 2.17 **EMSA reaction of the CBF3 core complex without λ protein phosphatase.** List of the components and the volumes used to make a 20 μL reaction of the CBF3 core complex with either 33 bp or 56 bp CDEIII DNA without λ protein phosphatase.

Component	**DNA : CBF3 Core Complex**				
	1:0	**0:4**	**1:1**	**1:2**	**1:4**
DNA	1.25 μM	–	1.25 μM	1.25 μM	1.25 μM
CBF3 Core Complex	–	5.00 μM	1.25 μM	2.50 μM	5.00 μM
10x Reaction Buffer	1x	1x	1x	1x	1x
$ZnCl_2$	10 μM	10 μM	10 μM	10 μM	10 μM
Glycerol	10%	10%	10%	10%	10%
10x PMP Buffer	1x	1x	1x	1x	1x
$MnCl_2$	1 mM	1 mM	1 mM	1 mM	1 mM
λ Protein Phosphatase	200 units	200 units	200 units	200 units	200 units
Total Volume	20 μL	20 μL	20 μL	20 μL	20 μL

Table 2.18 **EMSA reaction of the CBF3 core complex with λ protein phosphatase.** List of the components and the volumes used to make a 20 μL reaction of the CBF3 core complex with either 33 bp or 56 bp CDEIII DNA and λ protein phosphatase.

Ndc10 domains 1-2

The EMSA assay for Ndc10 domains 1-2 with fluorescently labelled 33 bp CDEIII and 56 bp CDEIII was assembled as described in Table 2.19. The assembled reactions were incubated at room temperature for 40 minutes. The EMSA assay was conducted,

imaged and the band shift was determined as described for the CBF3ΔN core complex EMSA assay

Component	DNA : Ndc10 domains 1-2				
	1:0	**0:4**	**1:1**	**1:2**	**1:4**
DNA	3.2 μM	–	3.2 μM	3.2 μM	3.2 μM
Ndc10 D1-2	–	12.8 μM	3.2 μM	6.4 μM	12.8 μM
10x Reaction Buffer	1x	1x	1x	1x	1x
Glycerol	10%	10%	10%	10%	10%
Total Volume	20 μL	20 μL	20 μL	20 μL	20 μL

Table 2.19 **EMSA reaction components for Ndc10 domains 1-2.** List of the components and the volumes used to make a 20 μL reaction for the EMSA assay of Ndc10 domains 1 and with 33 bp and 56 bp CDEIII.

2.9 Cryo Electron Microscopy theory

2.9.1 Introduction

The small wavelength of an electron enables high-resolution 3D imaging of biological material using cryogenic electron microscopy (Cryo-EM) (Egelman, 2016). Cryo-EM as a method for structural determination provides specific advantages which enables the user to overcome limitations of other methods such as being able to image large proteins and protein complexes, avoiding the need to crystallise the sample before imaging, which is the limiting step to structural determination for many samples and the ability to deal with structural heterogeneity (Cheng et al., 2015). Recent developments in detector technologies and image processing algorithms also have enabled us to obtain near atomic resolution for single particle protein or protein complex reconstructions (Kühlbrandt, 2014, Cheng et al., 2015).

2.9.2 Sample preparation

Prior to performing Cryo-EM, the sample must be prepared to withstand the microscope environment where the vacuum can cause dehydration and to also reduce the radiation

damage inflicted by inelastic scattering of electrons (J. Jensen, 2010). There are two common types of sample preparation: negative stain and vitrification.

Prior to imaging the sample using Cryo-EM the sample undergoes an initial screening using negative stain electron microscope (Negative stain EM). This is quality control measure to assess the appropriate sample concentration required for analysis, to assess the heterogeneity of the sample, the examine the dispersity of the sample and the size and shape of the sample of interest (Ohi et al., 2004, Cheng et al., 2015). The sample preparation involves applying a low concentration of a heavy salt, such as uranyl acetate, which surround the sample but are excluded from the volume occupied by the sample (Ohi et al., 2004, De Carlo and Stark, 2010). These stains readily interact with the electrons to produce strong contrast (Ohi et al., 2004, De Carlo and Stark, 2010). This method enables one to obtain a 3D reconstruction, providing the sample meets the quality control requirements discussed previously, however the grain size of the stain limits the resolution to approximately 18 Å (Burgess et al., 2004, AU - Scarff et al., 2018).

Vitrification

Cryo-EM is conducted on samples, which have been vitrified (Dubochet et al., 1988). The vitrified sample must have an ice thickness suitable to contain a high density of particles, in a number of different orientations with the thickness thin enough to achieve good contrast (Cheng et al., 2015). A number of considerations must be accounted for when preparing vitrified samples. First interactions between the sample and the air water interface can lead to preferred orientations of the sample on the grid or it can cause the sample to precipitate (Dubochet et al., 1988, Glaeser, 2018). This can be dealt with by using an increased ice thickness but this would reduce sample contrast (Orlova and Saibil, 2011). A support film can be used such as graphene oxide, which adsorbs the sample or the sample can be chemically fixed using a cross linking agent to ensure it remains intact (Stark, 2010, Russo and Passmore, 2016).

2.9.3 Image Formation

In microscopy the resolution obtainable is limited to half the wavelength of the illumination source. Electrons have a very small wavelength, which makes it a good source for imaging biological material. The electron microscope is composed of an electron emission source, typically a tungsten filament or a Field Emision Gun (FEG)

(Orlova and Saibil, 2011). The emission source is connected to a high voltage source and in tandem they emit electrons into a vacuum, where the accelerated electron beam is focused and directed to the sample by the electromagnetic lens system and subsequently processed by the detector (Orlova and Saibil, 2011, Cheng et al., 2015).

The emitted electrons are directed to the sample such that extensive radiation damage does not occur (Orlova and Saibil, 2011). When electrons are directed to the sample the electrons could be deflected by outer orbital electrons of the atoms of the sample, the electrons can be deflected at high angles when colliding or nearly colliding with the nuclei of the atoms of the sample, or the electrons can backscatter where it is deflected in the direction at which it came from (Orlova and Saibil, 2011). There are two types of electron scattering; elastic scattering is when the scattered electrons don't lose energy, and these provide high-resolution information (Orlova and Saibil, 2011). Inelastic scattering is where the electrons lose energy by transferring it to the sample and these electrons contribute to background noise (Orlova and Saibil, 2011).

The electromagnetic lens system

The electromagnetic lens system directs the electrons to the sample and then to the detector (Fig 2.5) (Orlova and Saibil, 2011). The microscope contains three types of lens systems: the condenser lens system, the objective lens system and the projector lens system (Orlova and Saibil, 2011). The condenser lens gathers the electrons and focuses these electrons on to the sample, which sits in between the condenser and objective lenses (Orlova and Saibil, 2011). The objective lens focuses the electrons scattered from the sample by converting diverging electrons into a parallel beam, and magnifies the sample (Orlova and Saibil, 2011). The objective lens also enhances sample contrast by preventing electrons, which are deflected at high angles from reaching the image plane (Orlova and Saibil, 2011). Finally the projector lens, magnifies further the image coming from the other lenses and projects the electron beam on to the detector (Orlova and Saibil, 2011). There are also apertures within the microscope which are small discs which exclude electrons further away from a fixed distance of from the optical axis (Orlova and Saibil, 2011).

The detectors

Previously data collection was carried out using either photographic film or charge coupled device (CCD) camera. Photographic films were optimised for different electron voltages, had very fine pixels and a large detection area (Orlova and Saibil, 2011). However, the process of chemical processing, drying and digitisation for computational

analysis is not convenient (Orlova and Saibil, 2011). The CDD cameras cannot withstand high-energy electrons and as a result the electrons must first be converted into photons and then re-converted to photoelectrons (Orlova and Saibil, 2011, McMullan et al., 2014). The electron conversion introduces noise to the images. Also CCD cameras perform worse than photographic film at higher voltages of 200-300 keV and only performs slightly better than photographic film at lower voltages of 80-100 keV (Orlova and Saibil, 2011, McMullan et al., 2014).

The new generation of electron microscopes are equipped with direct electron detectors (Bammes et al., 2012, McMullan et al., 2016). These detectors are designed to detect electrons directly without having to first convert the electrons into photons (Bammes et al., 2012, McMullan et al., 2016). The detectors are able to withstand radiation, have pixels, which are large enough such that only one pixel is excited by an electron at a time (Bammes et al., 2012, McMullan et al., 2016). Unlike the previous digital detectors, the direct electron detectors collect movies composed of a series of images taken in rapid succession, which enables alignment of the collected image frames in the movie to correct for both sample drift and beam induced motions (Li et al., 2013, Cheng et al., 2015).

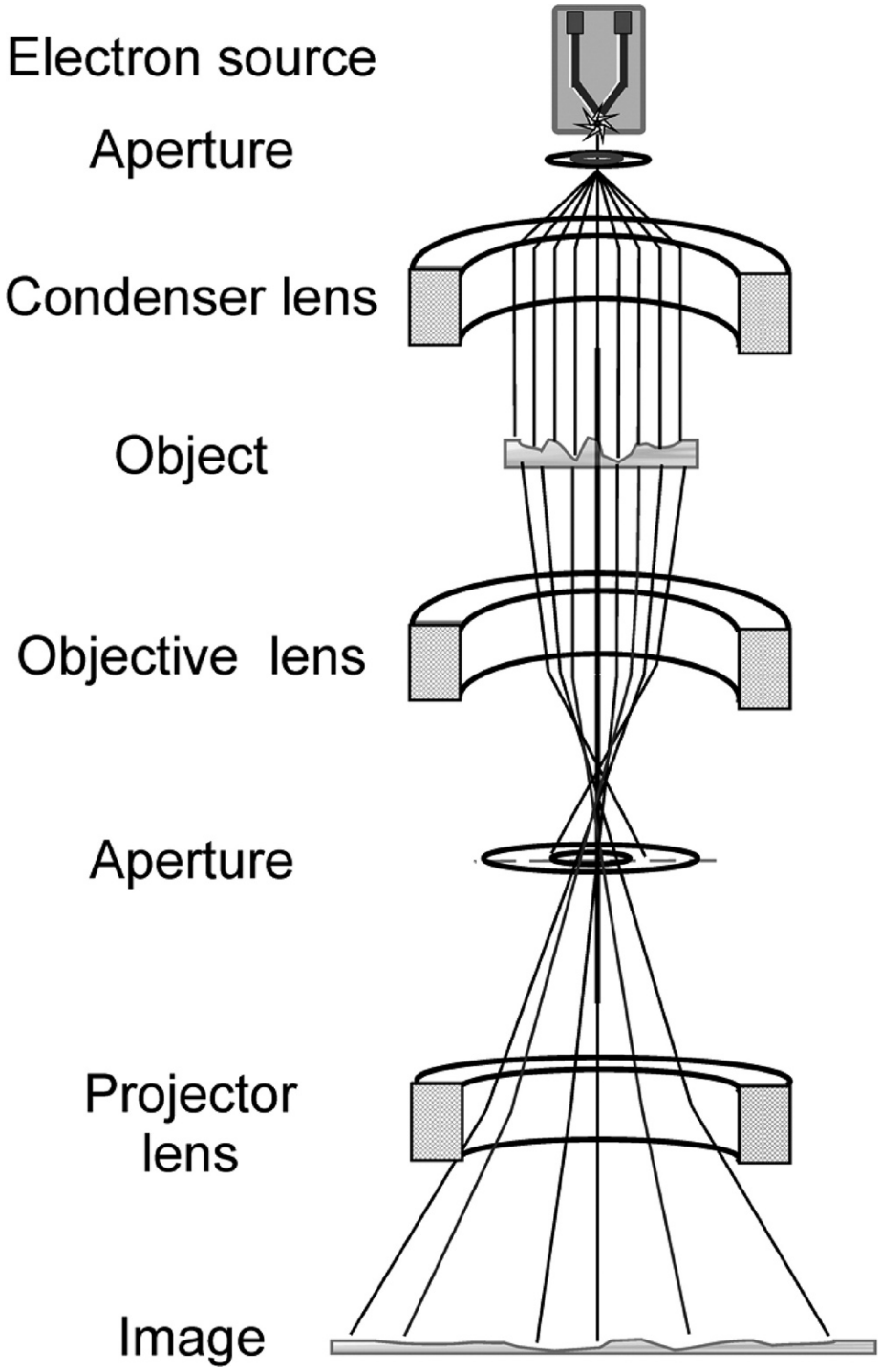


Fig. 2.5 **Schematic of a transmission electron microscope.** This figure illustrates how the electron source, the electromagnetic lenses and the aperture are organised. The diagram also describes the path of the electrons. Taken from (Orlova and Saibil, 2011)

2.9.4 Image contrast and Contrast Transfer Function

The contrast of an image is described mathematically as the magnitude variation between the brightest and darkest points of an image and the average intensity of across the whole image (Orlova and Saibil, 2011). Contrast arises from absorption of the incident electron beam. Transmission Electron Microscopy (TEM) utilises and modulates two forms of contrast: amplitude contrast and phase contrast (Wade, 1992, Orlova and Saibil, 2011). Biological samples are composed of light atoms which include hydrogen, carbon, nitrogen and oxygens, and as a result when the electron beam interacts with the sample, very few electrons are absorbed, with the majority of electrons scattered inelastically and elastically (Orlova and Saibil, 2011, Cheng et al., 2015). Amplitude contrast is obtained when the objective aperture excludes electrons scattered at a wide angle outside of a cut off limit, resulting in a change in the number of electrons in the exit wave relative to the entry wave (Orlova and Saibil, 2011).

Phase contrast results from the interference between scattered and unscattered electron on an image plane (Orlova and Saibil, 2011). The scattered electrons undergo a phase shift due to the scattering events and they undergo an additional phase shift because the electron has to traverse a greater path length relative to the unscattered electrons (Orlova and Saibil, 2011, Cheng et al., 2015). Interference between the scattered and unscattered electron beams can result in negative contrast, where the two waves are out of phase, positive contrast, where the two waves are in phase and no transfer of contrast when the scattered beam is 90°C out of phase relative to the unscattered beam (Orlova and Saibil, 2011). This is mathematically described in fourier space as the Contrast Transfer Function (CTF) (Wade, 1992). The interference between the two beams can be plotted as a sine function relative to the spatial frequency (Wade, 1992). In the CTF graph, there are multiple crossing at 0, which represents information that is lost (Orlova and Saibil, 2011). To compensate for the lost information, imaging is conducted across a defocus range for two reasons (Erickson et al., 1971, Orlova and Saibil, 2011). Imaging at defocus increases the path length of the scattered electrons to the detector relative to the unscattered electrons, which results in higher image contrast (Orlova and Saibil, 2011). By imaging at different defocus values, it enables one to obtain information across all spatial frequencies to compensate for the zero crossings for any particular defocus value (Fig 2.6) (Orlova and Saibil, 2011).

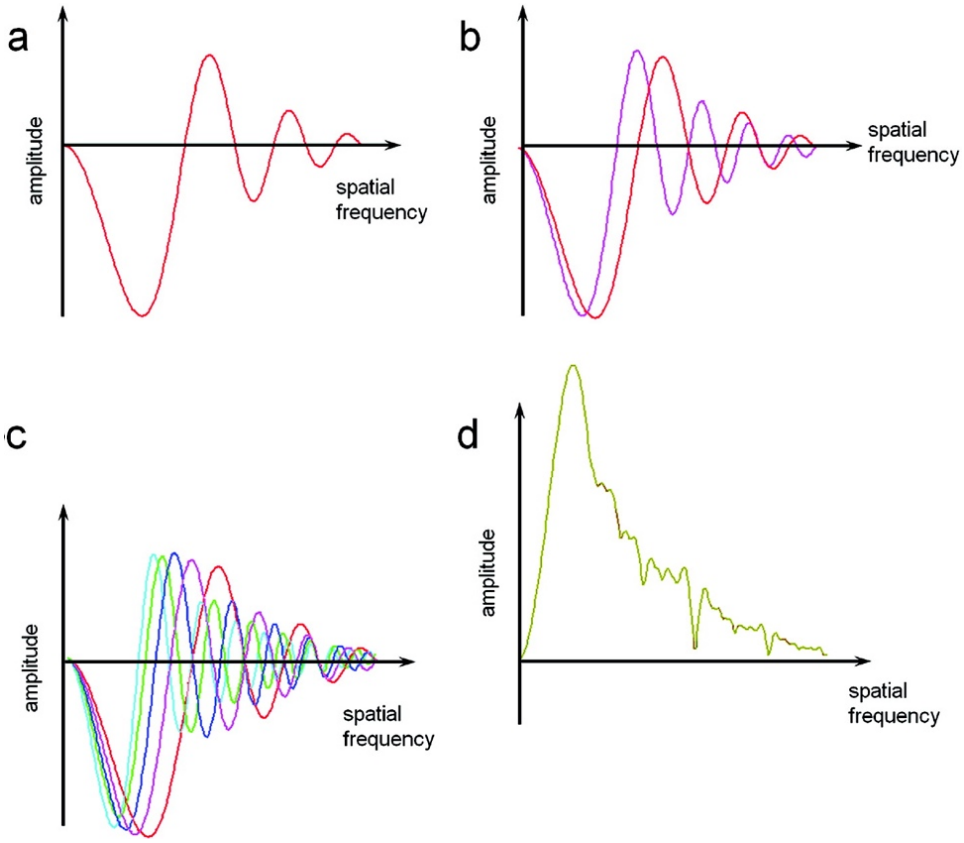


Fig. 2.6 **Contrast Transfer Function.** (A)CTF curve for a single defocus value (B)CTF curve for two defocus value (C) CTF curve for multiple defocus value (D) Sum of the amplitude values from the CTF curve for multiple defocus value from (C).Taken from Orlova and Saibil, 2011

2.9.5 Estimation and correction of Contrast Transfer Function

Estimation of CTF and correction is an essential part of image processing to ensure accurate image restoration and provides a method of quality control for assessing the quality of the collected micrographs. The necessity of CTF correction arises from distortions caused by microscope optics and operating conditions effect image appearance and contrast by effecting the signal transfer through the microscope, and without correction will result in local density displacement (Orlova and Saibil, 2011). Factors which can effect the images collected include but not limited to, lens aberrations, microscope instabilities and environmental factors such as temperature changes and

mechanical stability of the microscope (Cheng et al., 2015). The effect of these factors can be described in real space using the point spread function and in fourier space using the envelope function (Orlova and Saibil, 2011). CTF correction is performed on the micrographs separately, and enables the user to then combine the images together to perform image processing and analysis (Cheng et al., 2015).

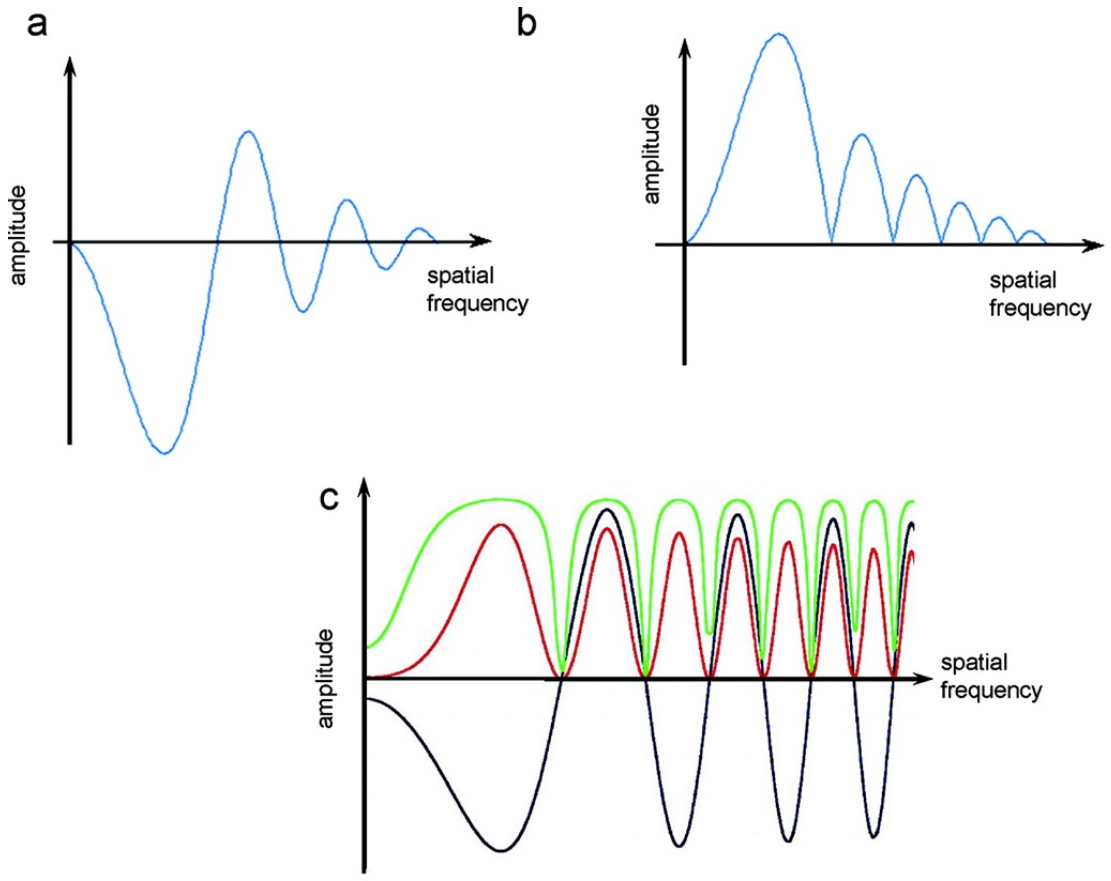


Fig. 2.7 **CTF correction.** (A)CTF curve from a single defocus, uncorrected (B)CTF curve after phase correction. (C) Comparison between the uncorrected CTF (black), rescaled (red) and amplitude corrected (green).Taken from (Orlova and Saibil, 2011)

To estimate CTF the following parameters are required: acceleration voltage, spherical aberration, defocus values, astigmatism and the percentage amplitude contrast (Mallick et al., 2005). The voltage and spherical aberration values are provided by the manufacturer. The percentage of amplitude contrast is estimated to be between 4 – 10% for cryo imaging, and the percentage used depends on the voltage the microscope

operates at (Toyoshima and Unwin, 1988, Orlova and Saibil, 2011). The defocus value is set during the data collection but this value is not accurate. To obtain accurate values for defocus and astigmatism, a calculated CTF is compared against the thon rings from the image power spectrum of the experimental data (Mindell and Grigorieff, 2003). There are two forms of CTF correction that are performed (Fig 2.7) (Penczek, 2010). Phase correction is performed to flip the phase of the negative CTF components (Orlova and Saibil, 2011). This allows for combination of the data to fill in the missing information from different defocus values at the zero crossing (Orlova and Saibil, 2011). Amplitude correction is performed to correct for the amplitude decay which occurs at high spatial frequencies which encode high resolution image information (Orlova and Saibil, 2011, Zhang, 2016).

2.9.6 Single particle 3D Image reconstruction

Motion Correction

To obtain higher resolution reconstruction, direct electron detectors collect images as movies to enable correction for beam induced motions and drift (Fig 2.8) (Li et al., 2013). The correction of these images is performed using MotionCorr or MotionCorr2 (Li et al., 2013, Zheng et al., 2017). MotionCorr aligns all the subframes from a single exposure, estimate the relative motion between these frames and then correct for whole frame image motion (Li et al., 2013). This has been shown to restore high resolution information in Fourier space (Li et al., 2013). This algorithm has been expanded in MotionCorr2 to correct for non-uniform local motion within each subframe (Zheng et al., 2017). The image is divided into patches, and the motion within each patch are iteratively determined and corrected (Zheng et al., 2017).

Particle picking

Particle picking is performed once pre-processing steps have been completed (Cheng et al., 2015). These pre-processing steps include motion correction and CTF estimation which is used to select micrographs for subsequent structure determination. Particle picking can be performed using a large number of different software's. Particle picking can be performed manually, semi-automatically, where the user performs manual picking on a small number of micrographs to provide a template for automatic picking or automatically (Zivanov et al., 2018). For particle picking, the expected diameter of the particle must be provided or a value for the box size for picking. The picking can

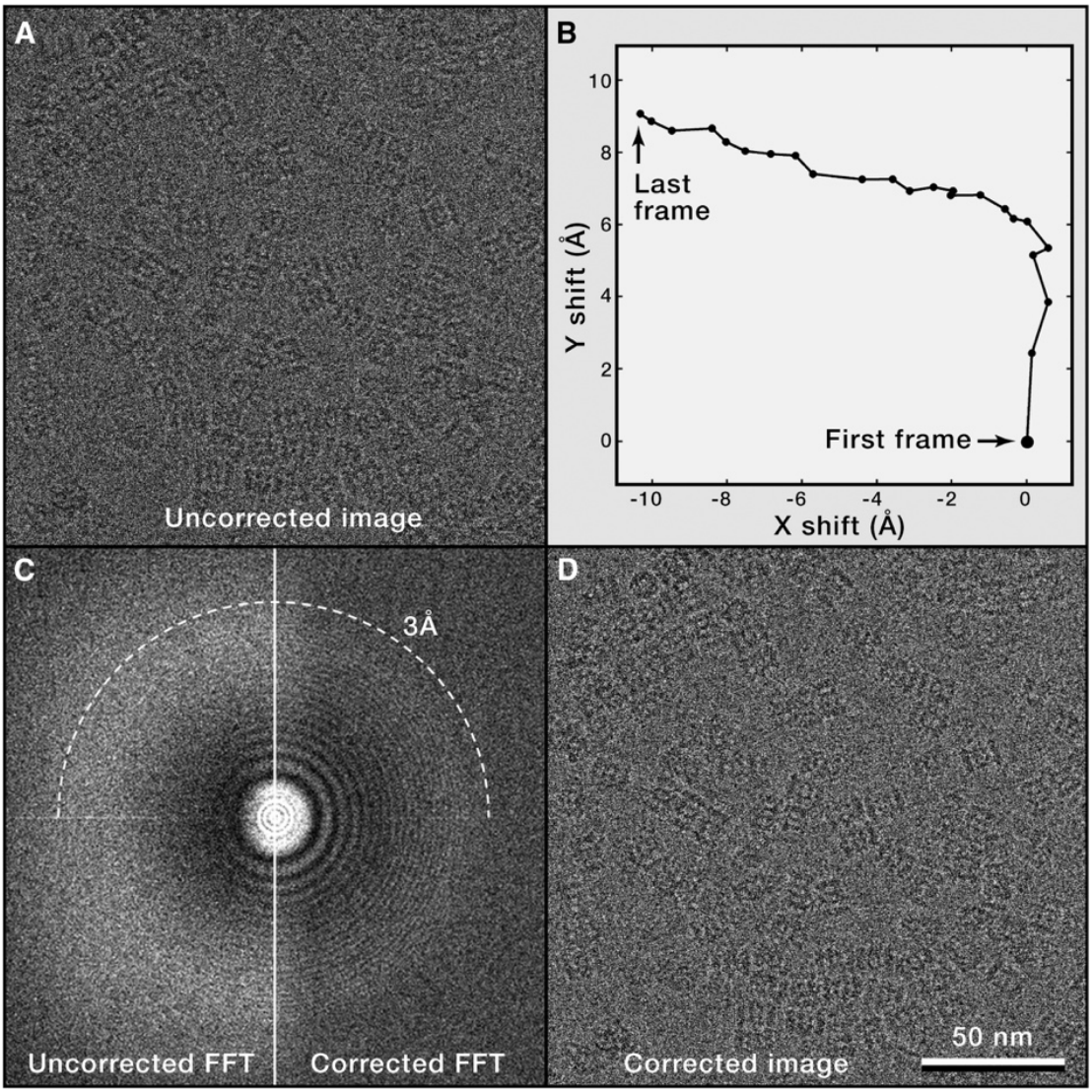


Fig. 2.8 **Motion correction.** (A)Non-motion corrected micrograph. (B)Graphic plot of motion across all the movie frames. (C) Comparison of power spectrum from the sum of all the frames with and without motion correction (D) Motion corrected micrograph. Taken from Cheng et al. 2015.

be performed using a template, but the template must be low pass filtered, to avoid the 'Einstein from noise' effect (Henderson, 2013).

2D Classification

Individual particles have low signal to noise ratio (SNR) which poses problems for alignment due to difficulty in assigning position and orientation of the particle (Orlova and Saibil, 2011). 2D classification separates the picked particles into a defined number of homogenous classes, boosting the SNR for easier alignment (Orlova and Saibil, 2011, Sigworth, 2015). 2D classification provides a method to help separate the particles of interest from artefacts, contaminants and empty fields, helps to separate the heterogeneity within the sample and provides a method for analysis of angular distribution to ensure there are enough particle views for efficient 3D reconstruction calculation (Cheng et al., 2015).

3D Classification and Refinement

3D classification provides a method of further improving the dataset, especially when working with heterogenous dataset, allowing the user to separate different conformations or different forms of a particular protein or protein complex (Cheng et al., 2015). To perform 3D classification, an initial 3D template is required called the initial model. The initial model can take the form of a low resolution 3D reconstruction from negative stain analysis of the sample or a low pass filtered X-ray crystal structure or EM map of the protein or protein complex (Cheng et al., 2015). Alternatively, the initial model can be generated from the dataset, for example commonly used processing programs Relion 3.0 and CryoSPARC V2, use a method called Stochastic Gradient Descent (SGD) to generate the initial model (Punjani et al., 2017, Zivanov et al., 2018).The SGD algorithm utilises a small and random subset of particle images to approximate a 3D structure which it updates over a series of iterations, based on the calculated gradient (Fig 2.9). (Punjani et al., 2017). The initial model is used to generate 2D projections which can be mapped to the 2D projections from the dataset for accurate Euler angle assignment, over a number of iterations until a stable solution is reached (Orlova and Saibil, 2011). Once the dataset has been divided into 3D classes, the user can select 1 or multiple classes to then perform refinement. Refinement involves using projection matching algorithms to improve orientation parameters and Euler angle assignments between the projections from the select 3D class and the dataset (Cheng et al., 2015).

Fourier Shell Correlation

Progression of refinement is measured and monitored using the 'gold-standard' Fourier Shell Correlation (FSC) curves (Henderson et al., 2012). The FSC curve is a plot of SNR as a function of spatial frequency and map resolution (Sousa and Grigorieff, 2007).

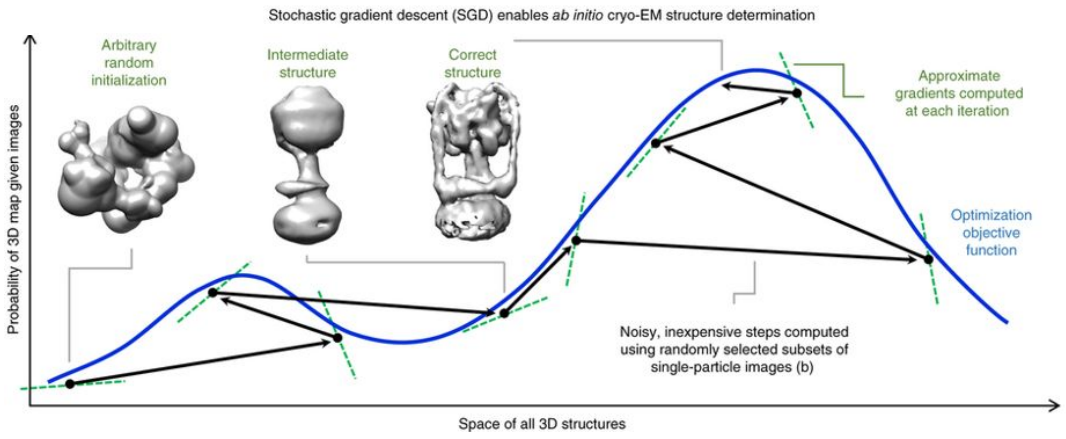


Fig. 2.9 **SGD initial model generation.** Graph showing the approximations of the 3D structure and the gradual improvement over iterations.Taken from (Punjani et al., 2017)

To generate an FSC curve, the dataset is divided into two at random to produce a 3D volume from each dataset (Sousa and Grigorieff, 2007). The Cross correlation of each model is then calculated in Fourier space and compared across different resolutions (Orlova and Saibil, 2011). The estimated resolution of the model is defined as the point at which the signal and noise equate to one another (Liao and Frank, 2010). Typically, two cut offs are used, 0.143, which is based off of x-ray structural determination or the more conservative 0.5 (Liao and Frank, 2010).

2.10 Cryo Electron Microscopy Data collection

2.10.1 Data Collection

Negative Stain Data

Negative stain cryo-EM was performed by Dr Natasha Lukoyanova, to screen micrographs prior to cryo-EM data collection. The data was collected on the Tecnai F20 microscope (FEI) at 200 kV with a pixel size of 0.92 Å/pix with a defocus range of -0.4 to -1.2 μM. The pixel size was very small compared to the 3-5 Å/pix typically used for

negative staining and due to the particle congregating close to one another, resulting in particle overlap.

CryoEM Data
The data collection was performed by Dr Natasha Lukoyanova. 0.13 mg/mL of sample was applied to glow-discharged Cu R1.2/R1.3 holey-carbon grids (Quantifoil Micro Tools) with graphene oxide support. The data was collected on an FEI Titan Krios at 300 kV with an electron dose of 51.4 e/Å^2 using K2 detector on counting mode. The pixel size used was 0.85 Å/pix with a defocus range of -1.5 μM to -3.6 μM. The images were collected at magnification of 165,000, with 2330 images collected.

2.10.2 Image processing

Motion correction
Motion correction was performed using Relion 3.0 MotionCorr2 implmentation (Zheng et al., 2017). The alignment and correction of the micrographs was performed on a 5 x 5 grid. Dose-weighted and non-dose weighted images were generated from motion correction. The non-dose weighted images were used for CTF estimation and subsequent particle picking and image processing was done on dose-weighted images.

CTF estimation and correction
CTF estimation was conducted on non-dose weighted micrographs using Gctf (Zhang, 2016). The estimated power spectra was used to remove micrographs with ice contamination, and poor CTF spectra. CTF correction was subsequently performed during 2D classification and 3D classification.

Particle picking
Particle picking was performed using Gautomatch v0.56 (developed by K. Zhang, MRC LMB, www.mrc-lmb.cam.ac.uk/kzhang). Unless otherwise stated, the picking was performed without templates with a specified particle diameter. The program then generates a box for picking based on the particle diameter. The particle coordinates were imported into Relion-3.0 and extracted using a specified box size without binning (Zivanov et al., 2018).

2D classification
2D classification was performed using Relion 3.0 (Zivanov et al., 2018). CTF correction was performed in 2D classification. The dataset was split into specified number of

classes and performed with 25 iterations. Mask diameter was specified based on the particle size. 2D classification was performed with image alignment.

Initial model

Two methods were used to generate an initial model for 3D classification. The first was to generate the model using Relion 3.0 (Zivanov et al., 2018). Relion uses Stochastic gradient descent algorithm to generate the model (Punjani et al., 2017, Zivanov et al., 2018). The model is generated using the selected 2D classes, performing CTF correction. The dataset is contained in a single class, and a mask diameter was specified based on the particle size. The model was processed with C1 symmetry. The model was generated with 50 initial iterations, 200 in between iterations and 50 final iterations, with an initial resolution cut-off of 35 Å and a final resolution cut off of 15 Å.

The second method was to use the pdb2mrc function in EMAN2, to take a pdb model, adjust the pixel size and box size to match the dataset, and this model was then low-pass filtered to 60 Å using the relion image handler script (Tang et al., 2007, Zivanov et al., 2018).

3D classification

The initial model specified was low pass filtered to 60 Å and the classification was performed with C1 symmetry. The selected 2D classes containing the particles were further CTF corrected and divided into specified number of classes. The classification was performed across 25 iterations with a specified mask diameter.

3D image refinement

The best 3D classes were selected using subset selection in Relion 3.0 (Zivanov et al., 2018). The appropriate 3D volume from the 3D classification was used as a reference map, which was subsequently low pass filtered to 60 Å. The refinement was performed using C1 symmetry. CTF correction was performed and a appropriate mask diameter was specified based on particle size. The resolution was approximated using 'gold standard' Fourier Shell Correlation (FSC) (Henderson et al., 2012).

Map visualisation was performed with Chimera (Pettersen et al., 2004).

Chapter 3

Results 1: Purification of the CBF3 complex and its subunits

3.1 Introduction

A major hurdle in structural analysis of the CBF3 complex is the ability to purify the protein complex at a suitable concentration for structural analysis. Prior to commencing this project, structural analysis of the CBF3 complex was limited to the individual subunits, including 2 truncated Cep3 crystal structures, 2 truncated Ndc10 structures and crystal structures of Skp1 in complex with a number of different F-box proteins (Bellizzi et al., 2007, Purvis and Singleton, 2008, Cho and Harrison, 2012, Perriches and Singleton, 2012). In order to do more comprehensive structural analysis and biochemical experiments, one must establish a methodology to purify the protein complex at a reasonable protein concentration with very little contamination from other proteins, which may non-specifically associate with the proteins or the individual chromatography columns.

This chapter will address the aim of establishing methodologies to express and purify both the wild type and the truncated CBF3 complexes in budding yeast. The lab had already established a method for expression and purification of the truncated CBF3 core complex (CBF3ΔN core complex) where the Zn_2Cys_6 domain of Cep3 was truncated (Zhang et al., 2018). This is due to the instability of the core complex, making the purification not feasible at the time. The lab also established a method to reconstitute

the truncated CBF3 complex where the Zn_2Cys_6 domain of Cep3 was truncated and the Ndc10 domains 3 to 5 were truncated (Zhang et al., 2018). The reconstitution was necessary due to the complex not remaining intact during the purification.

This chapter describes the development of a method for expression and purification of the full-length core CBF3 complex and a method to purify a co-expressed and intact truncated CBF3 complex where the Zn_2Cys_6 domain of Cep3 was truncated and the Ndc10 domains 4 and 5 were truncated. Furthermore, previous work on Ndc10 was conducted using bacterial expression systems. This chapter also describes the development of a methodology to express Ndc10 domains 1 to 2 and Ndc10 domains 1-3 in *S. cerevisiae* cells and the subsequent purification.

3.2 Protein expression of Ndc10

Typically when expressing the truncated CBF3 core complex in *S. cerevisiae*, across 24 litres of cell culture, using the previously established protocol from the lab (Section 2.4.3), one would expect to get a cell yield of approximately 240 g. When inducing protein expression using galactose, the cells would take 8 hours to reach optical density. When the core complex is co-expressed with Ndc10 or Ndc10 domains 1 to 2 or Ndc10 domains 1-3, the cell yield reduces to 150 g and the cells grow slower. This results in a smaller yield of protein or protein complex.

It was hypothesised that the doubling of the time taken to reach optical density and the reduced cell yield could be a result of leaky expression of Ndc10 causing toxicity or providing unnecessary burden on the cells. To test this, 0.1% glucose was added to the media, which represses the promoter. The 0.1% was chosen as its enough glucose to repress leaky expression but sufficiently small concentration of glucose, which can be outcompeted, by the 2% of galactose needed for protein induction. By inhibiting the leaky expression, time taken to reach optical density was reduced to approximately 7 hours and a cell yield of approximately 240 g was routinely obtained from 24 litres of protein expression.

3.3 Protein expression of Core CBF3 Complex

The lab had established a protocol to express the CBF3Δ core complex (Section 2.5.1). This protocol was sufficient to express the full length CBF3 core complex, with two minor adjustments made. First 100 μM $ZnCl_2$ was added to the media prior to protein induction with galactose, to ensure sufficient zinc is available for binding and stabilising the Zn_2Cys_6 cluster domain of Cep3. The second change was made to the lysis buffer, which was adjusted to suit the chromatography steps used to purify the full length complex.

3.4 Protein Purification of the CBF3 core complex

The CBF3ΔN core complex was very stable and the subsequent structure determination of the complex revealed, high levels of detail relating to the Ctf13 protein, as well as how the subunits are arranged relative to one another and identified a potential DNA binding channel (Zhang et al., 2018). However, detail related to the Zn_2Cys_6 domain were missing, which plays an important role in centromere recognition and binding (Purvis and Singleton, 2008). As a result, a method to purify the protein complex with intact zinc cluster domains is necessary to expand the understanding of how the complex functions. The protocol used to purify the CBF3ΔN core complex is not suitable for the full-length version, because calmodulin affinity purification requires the use of EGTA to elute the protein. If used in the presence of the zinc cluster domain, EGTA would strip the zinc atoms from the domain, rendering it unstable and possibly non-functional (NYBORG and PEERSEN, 2004).

The initial strategy was to adapt the purification protocol for the CBF3ΔN core complex and modify it by omitting the Calmodulin affinity purification and introducing Heparin chromatography. Constructs 1 and 3 (Table 3.1) were co-expressed in the *S. cerevisiae* expression system and lysed (Section 2.4.3 and 2.5.3). The lysate was cleared using ultracentrifugation at 38,000 rpm for 1 hour 30 minutes at 4°C. The lysate was incubated with TALON® Cobalt resin (Clontech). The resin was washed with buffer A (20 mM Tris-Cl, 500 mM NaCl, 20 mM Imidazole, 10 μM $ZnCl_2$, 5 mM β-mercaptoethanol) and the protein was eluted in buffer B (20 mM Tris-Cl, 500 mM NaCl, 500 mM Imidazole, 10 μM $ZnCl_2$, 5 mM β-mercaptoethanol). The sample was subsequently concentrated, the salt concentration was diluted 150 mM using QA buffer

Construct	**Proteins**	**Vector and selection marker**	**Protein tag**
1	Cep3 Skp1	pRS-426 TRP	Cep3 (N-term 6x His)
2	Sgt1 Ctf13	pRS-424 URA	Sgt1 (C-term Strep II) Ctf13 (C-term CBP)
3	Sgt1 Ctf13	pRS-424 URA	Sgt1 (C-term Strep II) Ctf13 (N-term 6x His)
4	Cep3 Skp1	pRS-426 TRP	No tags
5	Sgt1 Ctf13	pRS-424 URA	Sgt1 (C-term Strep II) Ctf13 (C-term 6x His)
6	Sgt1 Ctf13	pRS-424 URA	Sgt1 (C-term Strep II) Ctf13 (C-term Strep II)

Table 3.1 **CBF3 core complex plasmids.** Plasmids constructs used to purify the full-length CBF3 core complex. These constructs were gifted by Dr. Wenjuan Zhang and Jonathan Lucas in the Vaughan lab.

(20 mM Tris-Cl pH8, 100 mM NaCl, 10 % Glycerol, 10 μM $ZnCl_2$, 2 mM DTT), loaded on to 1 mL Resource^{TM} Q (GE Healthcare) anion exchange column and eluted with buffer QB (20 mM Tris-Cl pH 8, 1 M NaCl, 10 % Glycerol, 2 mM DTT, 10 μM $ZnCl_2$) across a linear gradient. The sample was then concentrated and diluted to 150 mM salt concentration and subsequently loaded on to a 1 mL HiTrap® Heparin HP (GE Healthcare). The sample was eluted in the QB buffer used for 1 mL Resource^{TM} Q. The sample was then incubated with 33 bp CDEIII DNA for 12 hours at 4°C. The sample was then loaded on to Superdex 200 10/300 GL (GE Healthcare) where the sample was eluted into the buffer: 15 mM Tris-Cl pH 8, 200 mM NaCl, 10 μM $ZnCl_2$, 2 mM DTT.

This method did not yield an stable CBF3 core complex because the protein complex precipitated at multiple stages of the procedure (Fig 3.1). First the complex precipitated post elution from the TALON® resin, whilst on ice. The protein also precipitated whilst attempting to concentrate the sample prior to the anion exchange chromatography. The protein complex proceeded to precipitate further when the salt concentration was diluted to 150 mM. This was the step at which most of the protein complex tended to be lost. This is reflected in the significant decrease in SDS-PAGE band intensity continuously between each step of the purification (Fig 3.2). Prior to loading the sample on to the gel filtration column, the sample is centrifuged at 13,200 rpm for 10 minutes. The precipitate from this was also loading on the SDS-PAGE, which shows bands corresponding to Cep3, Ctf13 and Skp1. The small yield of protein is highlighted by the negative stain-EM of the sample. The data shows very few particles which are not homogenous in size. From these results it was inferred either prolonged exposure to imidazole or the lowering of salt concentration caused the instability.

The second strategy centred on cleaving the complex directly from the TALON® resin which removes the need for imidazole. This strategy used constructs 1 and 3 (Table 3.1). Thrombin protease was used to cleave at the thrombin cleavage site upstream of the N-terminal His tag of Cep3 in construct 1 (Table 3.1). The expression, lysis and clearing of the lysate was the same as the first strategy, using construct 1 and 2.

The cleared lysate was incubated with TALON® resin (Clontech) for 1 hour at 4°C. The resin was washed with 100 mL of Buffer (20 mM Tris-Cl pH 8, 500 mM NaCl, 20 mM Imidazole, 10 mM β-mercaptoethanol, 10 μM $ZnCl_2$, 10 % glycerol). The resin was resuspended in the same buffer and 1 unit of thrombin (GE Healthcare) was added to the resin, for 12 hours at 4°C. After the incubation, a visible precipitate formed on top

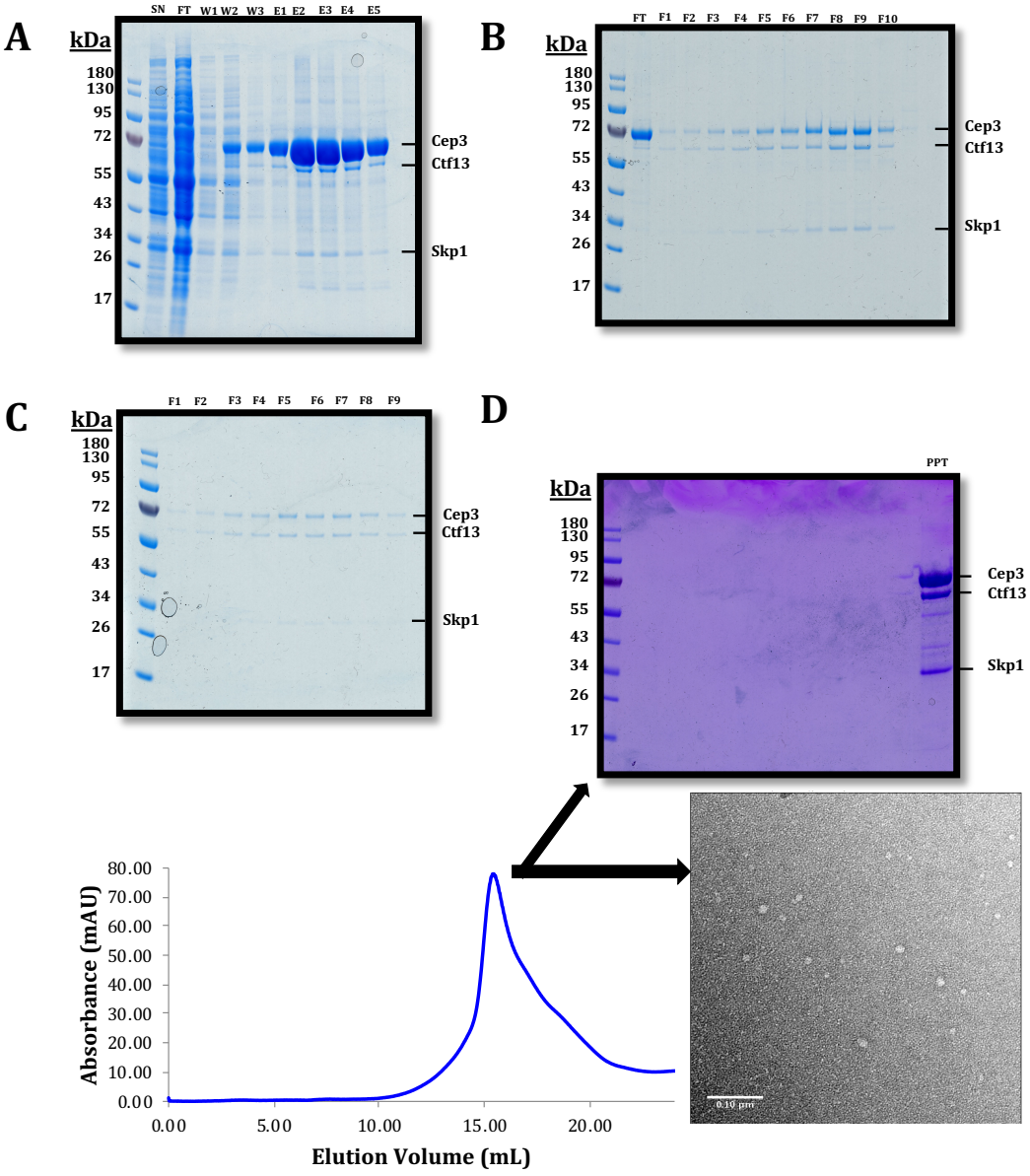


Fig. 3.1 **Purification of CBF3 core complex - strategy 1.** (A) SDS-PAGE post cobalt affinity chromatography. SN represents the supernatant,FT represents the flow through, W1 is wash 1 and E1 is elution1. (B) SDS-PAGE after anion exchange chromatography. The F1 is fraction 1 (C) SDS-PAGE after heparin chromatography. (D) SDS-PAGE of gel filtrated sample (PPT represents the precipitate), the chromatogram from gel filtration and the micrograph with a scale bar of 100 nm, of the gel filtrated sample.

of the solution. The precipitate was resuspended in 1 mL of 3x reducing SDS loading dye and was analysed on SDS-PAGE, which showed the subunits of the complex were present in the precipitate (Fig. 3.2). The rest of the sample was concentrated, the NaCl concentration was diluted to 150 mM and loaded on to a 1 mL Resource^{TM} Q (GE Healthcare). The column was washed with QA buffer (20 mM Tris-Cl pH 8, 100 mM NaCl, 10 % Glycerol, 10 μM $ZnCl_2$, 2 mM DTT) and eluted with buffer QB (20 mM Tris-Cl pH 8, 1 M NaCl, 10 % Glycerol, 2 mM DTT, 10 μM $ZnCl_2$) across a linear gradient. The amount of complex obtained from the ion exchange was deemed to be too small for gel filtration, a prerequisite for single particle analysis. Therefore, attempts to improve thrombin cleavage were pursued.

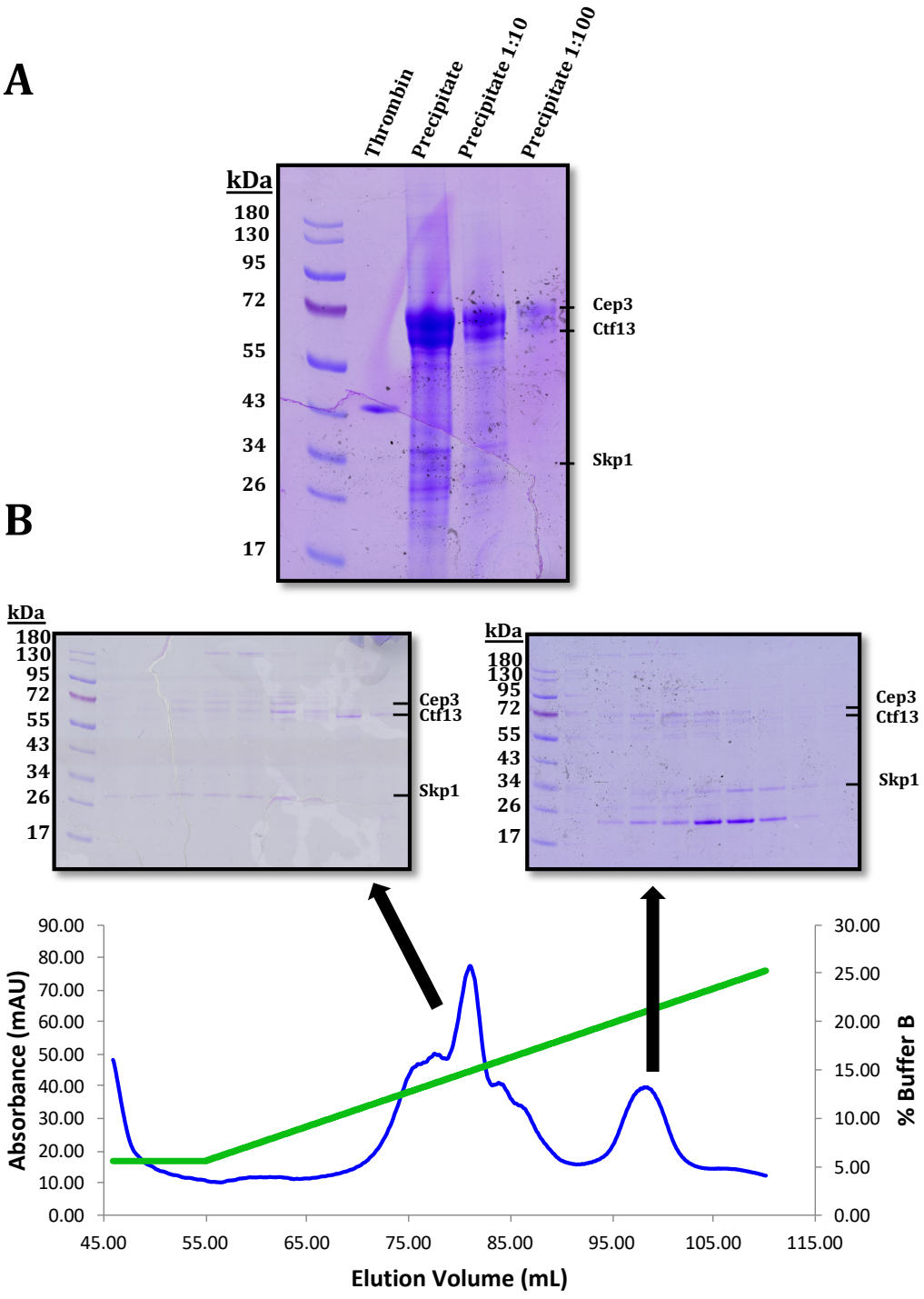


Fig. 3.2 **Cleavage of CBF3 Core complex - strategy 2.** (A) SDS-PAGE analysis of the precipitate after 12 hour incubation at 4 °C. (B) SDS-PAGE and the chromatogram after Ion exchange chromatography.

The protocol was modified to reduce the thrombin incubation time to 2 hours at 4°C , where the SDS-PAGE analysis showed a very small quantity of protein present in the cleaved fraction. The cleaved fraction was then applied to HiLoad 16/60 Superdex 200 PG (GE Healthcare) for gel filtration, avoiding ion exchange chromatography which resulted in significant protein precipitation, and eluted in gel filtration buffer (15 mM Tris-Cl pH 8, 200 mM NaCl, 10 μM $ZnCl_2$, and 2 mM DTT). However no protein could be visualised by SDS-PAGE (Fig 3.3). The protocol was also modified to conduct the thrombin cleavage at room temperature for 2 hours, for efficient cleavage reaction. This also led to precipitation (Fig 3.3.C).

The next strategy was to remove the His tag from Cep3. From co-expression experiments done previously in the lab, it was evident that Cep3 expresses at a higher level than Ctf13. Both Cep3 and Ctf13 have His tags (Constructs 1,3 and 5) (Table 3.1). If Cep3 expresses at much higher levels, then it was plausible that Cep3 binding would outcompete the complex. The purification strategy was centred on the purification strategy for Cep3 (Section 2. 6.10). Constructs 3 and 4 (Table 3.1) were co-expressed, the cells were lysed and the lysate was cleared as done previously. The cleared lysate was incubated with 5 mL TALON® Cobalt resin (Clontech) for 12 hours. The resin was washed with 100 mL of buffer A (20 mM Tris-Cl, 500 mM NaCl, 20mM Imidazole, 10 μM $ZnCl_2$, 5 mM β-mercaptoethanol and 10 % glycerol) and the protein was eluted in Buffer B (20 mM Tris-Cl, 500 mM NaCl, 500 mM Imidazole, 10 μM $ZnCl_2$, 5 mM β –mercaptoethanol and 10 % glycerol). The eluted protein was pooled, concentrated with an Amicon® Ultra Centrifugal filter (Merck), and applied to HiLoad 16/60 Superdex 200 PG (GE Healthcare) for gel filtration. The sample was eluted in buffer composed of 15 mM Tris-Cl pH 8, 200 mM NaCl, 10 μM $ZnCl_2$, and 2 mM DTT. The SDS-PAGE of the TALON® eluted fractions didn't show bands corresponding to the complex (Fig 3.4), as was the case for the SDS-PAGE of the gel filtrated sample. The SDS-PAGE of the cleared lysate is inconclusive when determining the presence of Cep3 and the lack of antibody against Cep3 also limits our ability to detect Cep3 expression, but the presence of Skp1 is evident.

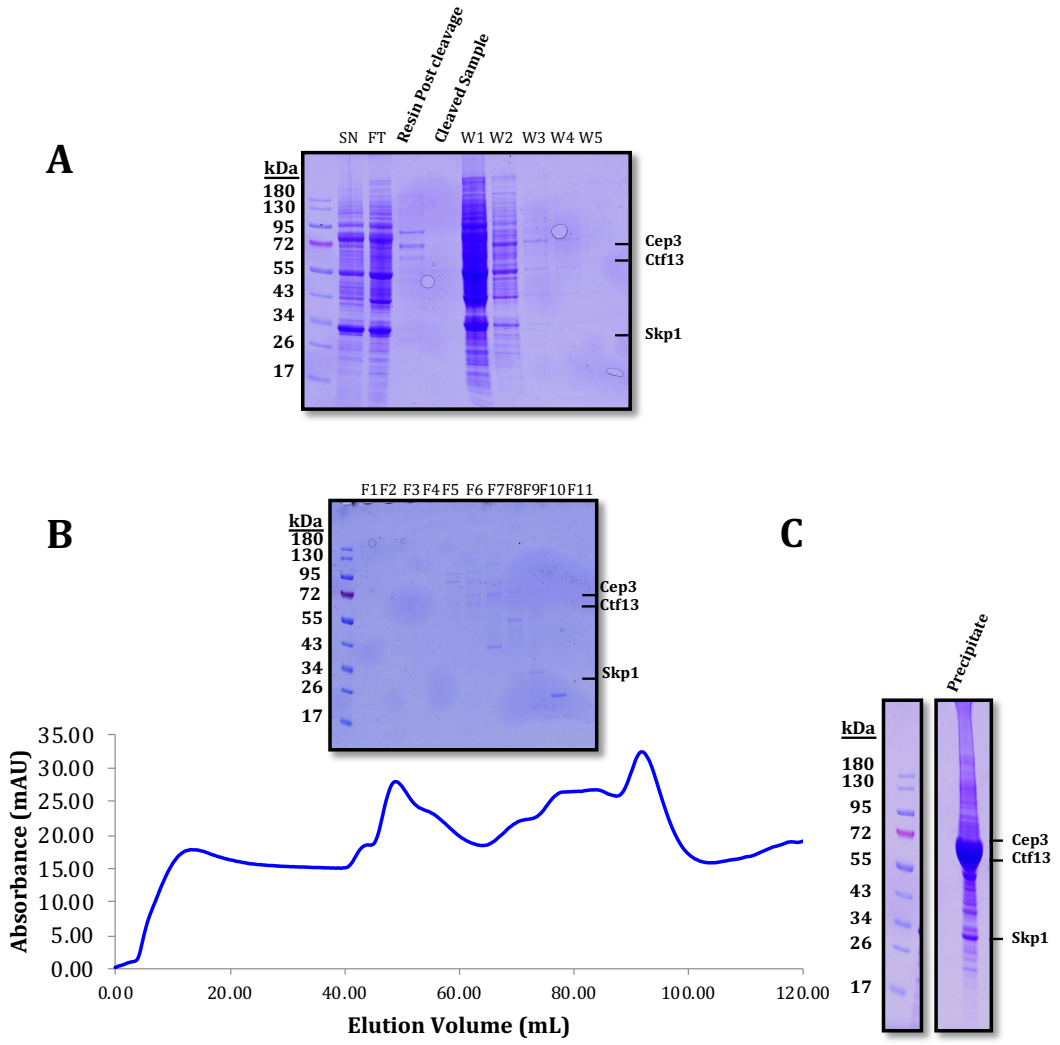


Fig. 3.3 **Cleavage of CBF3 Core complex - strategy 3.** (A) SDS-PAGE analysis of the cleaved sample after 2 hour incubation at 4°C. (B) SDS-PAGE and the chromatogram after gel filtration of the cleaved sample. (C) SDS-PAGE of the precipitate after 2 hour incubation at room temperature.

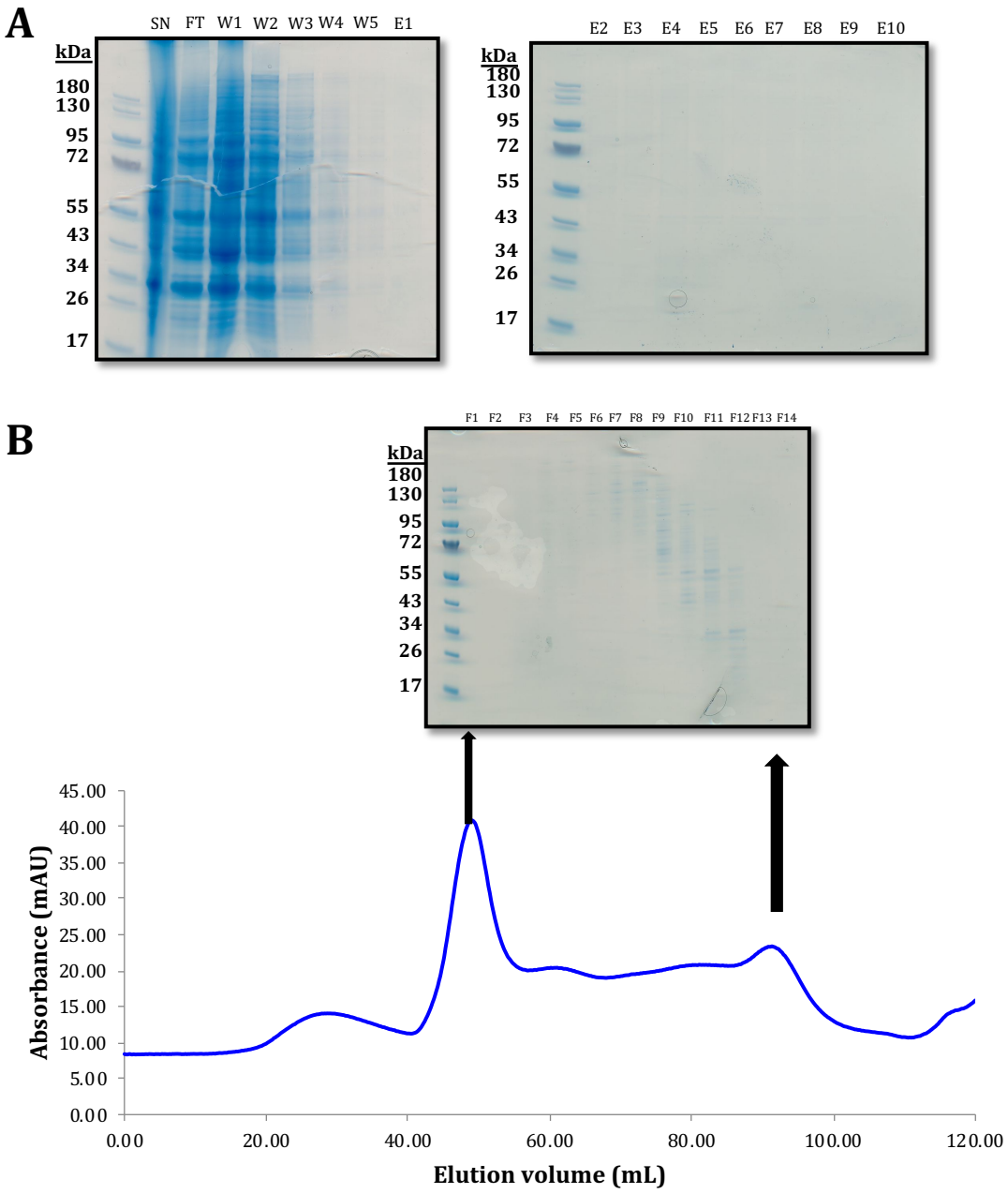


Fig. 3.4 **Cleavage of CBF3 Core complex - strategy 4.** (A) SDS-PAGE of the IMAC purification step, across two gels analysing the contents of the cleared lysate (SN), the unbound sample (flowthrough), the wash fractions (W1) and the elutions (E1). (B) The chromatogram and SDS-PAGE of the gel filtrated sample.

The next strategy was centred on three considerations. The successful purification of the CBF3ΔN core complex utilised Ctf13 with a C-terminal tag. Whereas, each of the previous attempts, described above were conducted with N-terminally tagged Ctf13. It was considered that tagging at the N-terminus could possibly introduce mRNA instability or potential protein folding problems for this particular protein, reducing the yield of expressed Ctf13. The second consideration was related to the Cep3 and Ctf13 having the same type of tag, discussed above. Third consideration was to utilise the DNA binding function of the CBF3 core complex to separate it from the contaminants. The new strategy involved the use of constructs 1 and 6 (Table 3.1). Constructs 1 and 6 were co-expressed as done previously and the cells were re-suspended in the 2x lysis buffer (40 mM Tris-Cl pH 8, 500 mM NaCl, 25 % glycerol, 2 mM Magneisium acetate, 4 cOmplete^{TM} EDTA-free protease inhibitor cocktail (Roche), 20 μM $ZnCl_2$, 0.2 % IGEPAL, 40 mM NaF, 4 μg/mL DNaseI and 3 mM DTT). The sample was lysed and cleared as done previously. The protein was applied to a pre-equilibrated 5 mL StrepTrap® HP (GE Healthcare) to interact with the STREP tagged on Ctf13. The sample was washed in Buffer A (20 mM Tris-Cl pH 8, 500 mM NaCl, 2 mM DTT, 10 μM $ZnCl_2$) and eluted in Buffer B (20 mM Tris-Cl pH 8, 500 mM NaCl, 2 mM DTT, 10 μM $ZnCl_2$, 3 mM Desthiobiotin). The peak fraction was pooled, the NaCl concentration was adjusted to 200 mM and then loaded on to a 1 mL HiTRAP® Heparin HP (GE Healthcare), where the loaded sample was washed with 20 % Buffer A and then the protein was eluted across a linear gradient from 20 % Buffer B to 50 % Buffer B. Buffer A was composed of 20 mM Tris-Cl pH 8, 10 % glycerol, 2 mM DTT, 10 μM $ZnCl_2$. Buffer B was composed of 20 mM Tris-Cl pH 8, 1M NaCl, 10 % glycerol, 2 mM DTT, 10 μM $ZnCl_2$.

The chromatogram after STREP affinity purification shows a large symmetrical peak of 2000 mAU (Fig 3.5.A). SDS-PAGE analysis of the peak shows large intense bands corresponding to Sgt1. This was expected because Sgt1 is co-expressed with the protein complex and is C-terminally tagged with STREP II. The SDS-PAGE also shows a protein band present between the 72 and 55 kDa markers, where Ctf13 is expected as well as a band at the expected point for Skp1. There is no obvious band corresponding to Cep3, but previous work from the lab has shown the presence of Ctf13 is dependent on its association with Cep3. Heparin chromatography was conducted to separate Sgt 1 from the protein complex.

After Heparin chromatography (Fig 3.5.B), the chromatogram showed a single wide peak, which was analysed on SDS-PAGE. The SDS-PAGE shows clearly a band

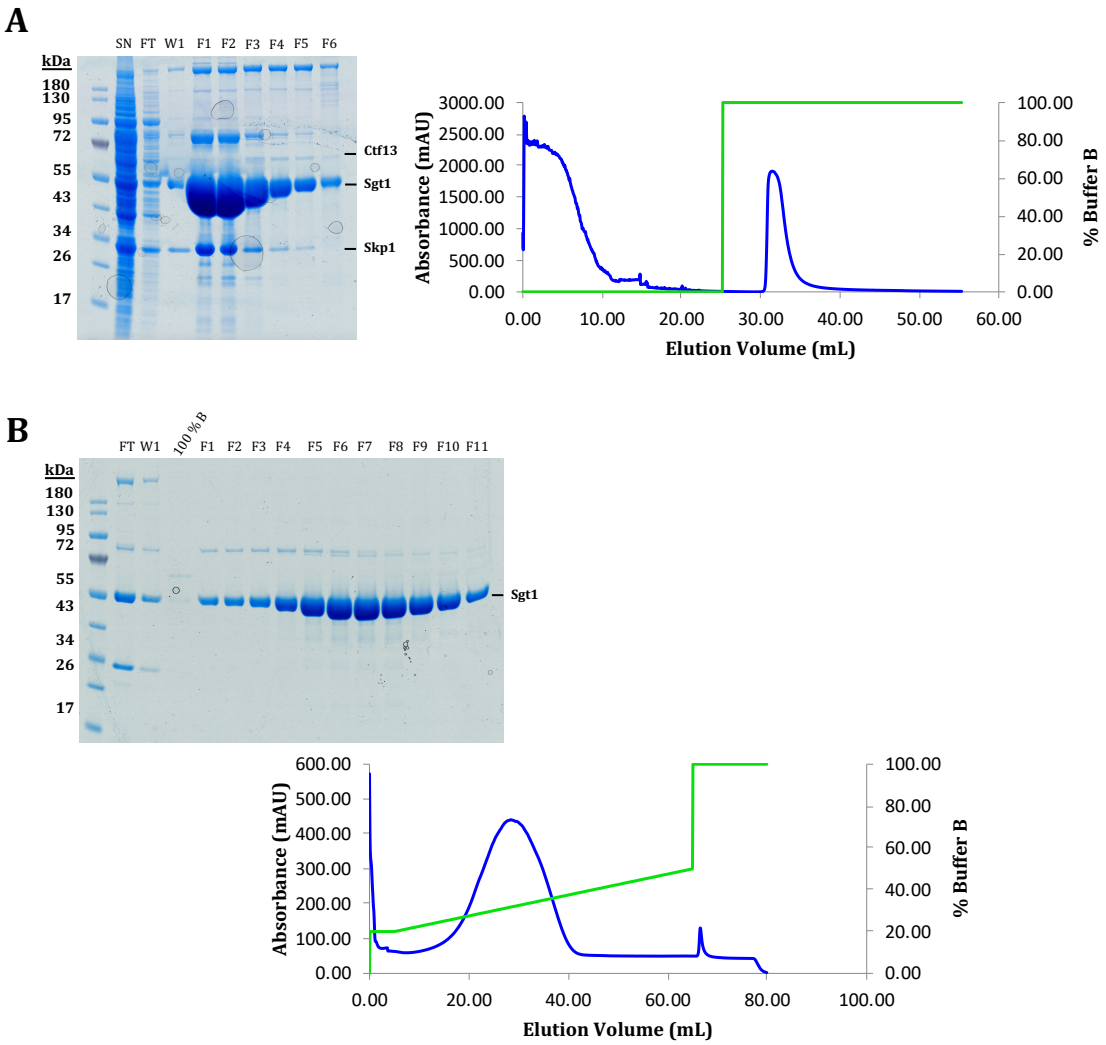


Fig. 3.5 **CBF3 Core complex with Strep tagged CBF3 - strategy 5.** (A) Chromatogram and SDS-PAGE after STREP affinity purification. (B) Chromatogram and SDS-PAGE after Heparin chromatography.

representing Sgt1 and no bands for Cep3, Skp1 or Ctf13. The STREP purification shows Sgt1 expresses in large quantity compared to CBF3 subunits, and likely oversaturates the column. Therefore, in order for this strategy to work, the STREP tag on Sgt1 had to be removed first before proceeding with this method.

The final strategy involved the use of His tags on both the Cep3 and Ctf13 as done previously, except the Ctf13 was tagged at the C-terminus this time. It was determined from previous strategies that at NaCl concentration below 200 mM, the protein

precipitates rapidly. The protein complex does not bind to the ion exchange resin at concentrations above 150 mM, but the complex does bind to Heparin resin at a NaCl concentration of 200 mM. Using the information gleaned from previous attempts a successful method was established, where constructs 1 and 5 (Table 3.1) were co-expressed as done previously. The cells were stored in a 2x lysis buffer composed of 40 mM Tris-Cl pH 8, 1 M NaCl, 25 % glycerol, 2 mM Magnesium acetate, 4 cOmplete^{TM} EDTA-free protease inhibitor cocktail (Roche), 20 μM $ZnCl_2$, 0.2 % NP-40, 100 mM Sodium fluoride, 4 μg/mL DNaseI and 5 mM β-mercaptoethanol.

The cells were lysed as described previously. After adjusting the pH the cells were sonicated for 3 cycles at an amplitude of 65 % for a total time of 2 minutes, with pulses on for 2 seconds and off for 4 seconds. The final cycle was carried out using the same settings at amplitude of 80 %. The sonication was carried out to shear the excess DNA. The lysate was cleared as done previously. After clearing the lysate, the sample was sonicated again for 3 cycles at an amplitude of 30 % for a total time of 1 minute, with pulses on for 2 seconds and off for 4 seconds. This was carried out in order to load the cleared lysate on to the chromatography columns. Without sonication, the sample blocks the column preventing the loading of the full sample volume.

The sample was first loaded on to a 5 mL HISTrap^{TM} with a Peristaltic Pump P-1 (GE Healthcare). Prior to loading, the column was equilibrated with 10 column volumes of lysis buffer. After loading the sample, the column was washed with Buffer A (20 mM Tris pH 8, 20 mM Imidazole pH 8, 200 mM NaCl, 10 μM $ZnCl_2$ and 5 mM β-mercaptoethanol) for 10 column volumes. Buffer B was composed of 20 mM Tris pH 8, 500 mM Imidazole pH 8, 500 mM NaCl, 10 μM ZnCl2 and 5 mM β-mercaptoethanol. After washing the column with Buffer A, the resin was washed with 8 % Buffer B for 10 column volumes to remove additional contaminants. The protein was eluted at 30 % Buffer B into a fraction plate containing imidazole dilution buffer (20 mM Tris pH 8, 200 mM NaCl, 10 μM $ZnCl_2$ and 5 mM β-mercaptoethanol) to reduce the sample imidazole concentration to 100 mM. The eluted fractions were analysed using SDS-PAGE. The fractions containing the complex were selected, pooled and the NaCl concentration was adjusted to 200 mM using salt dilution buffer (20 mM Tris-Cl pH 8, 10 % glycerol, 10 μM $ZnCl_2$ and 5 mM β-mercaptoethanol). The sample was then loaded on to a 1 mL HiTrap® Heparin HP (GE Healthcare) using a Peristaltic Pump P-1 (GE Healthcare), which was pre-equilibrated with 5 column volumes of Buffer HA (20 mM Tris pH 8, 10 % glycerol, 2 mM DTT, 10 μM $ZnCl_2$), 10 column volumes of Buffer HB (20 mM Tris pH 8, 10 % glycerol, 2 mM DTT, 10 μM $ZnCl_2$ and 1M NaCl),

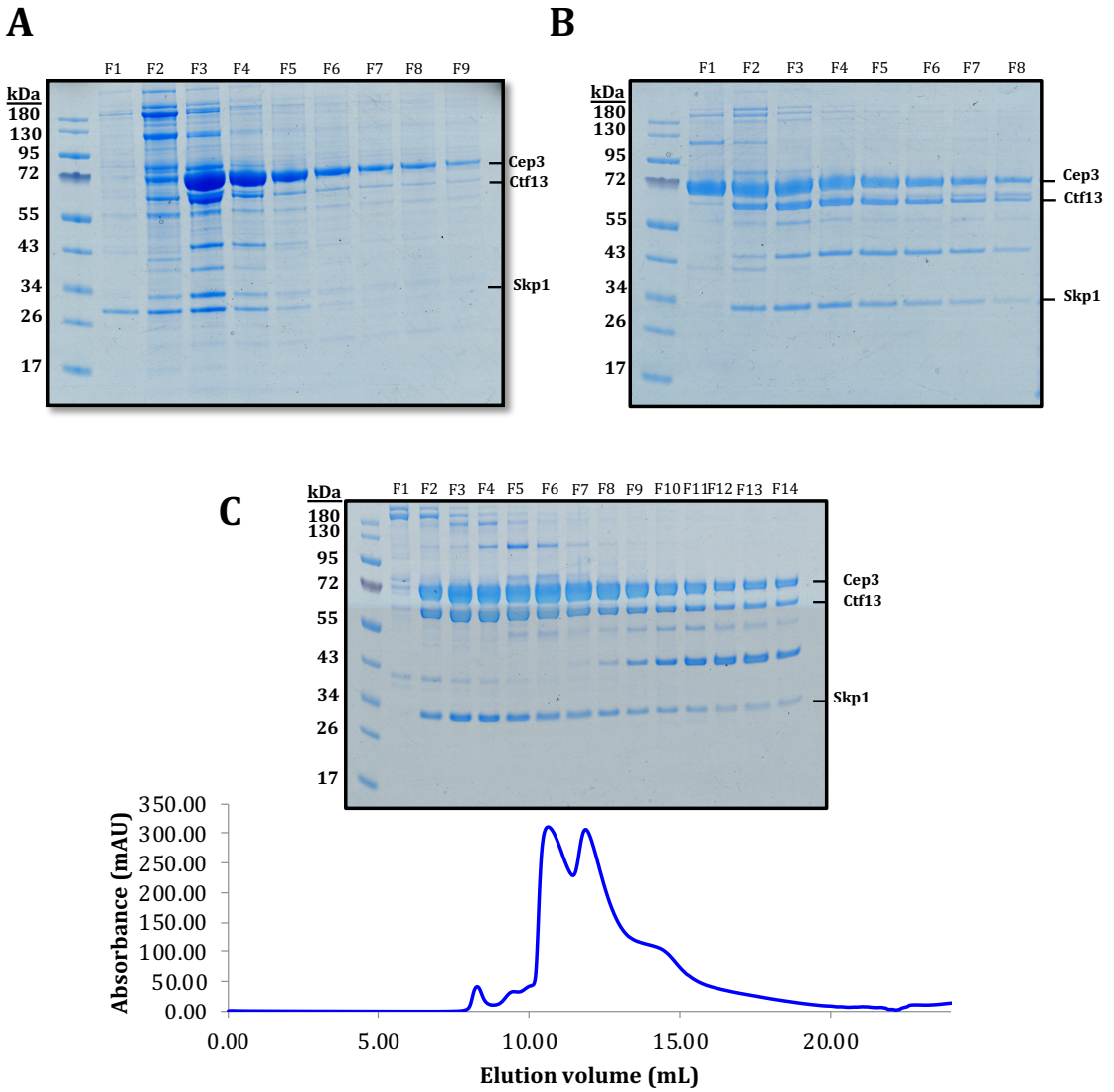


Fig. 3.6 **CBF3 Core complex purification - strategy 6.** Purification of the CBF3 Core complex. (A) SDS-PAGE after IMAC (B) SDS-PAGE after Heparin chromatography. (C) Chromatogram and SDS-PAGE after gel filtration.

followed by 15 column volumes of 20 % Buffer HB, which enables the salt concentration in the column to match the salt concentration of the sample. The resin containing the bound protein was washed with 20 % Buffer HB, to remove contaminants, then washed with 25 % Buffer HB to remove excess Cep3. The sample was eluted at 40 % Buffer HB. The eluted fractions were analysed using SDS-PAGE and the eluted fractions were concentrated with an Amicon® Ultra 15 mL Centrifugal Filters (Merck) and the concentrated sample was injected on to a Superdex increase 200 10/300 GL

Construct	Protein	Vector and selection marker	Protein tag
1	Ndc10 Domains 1 to 2	pRS-426	C-term Strep II
2	Ndc10 Domains 1-3	pRS-426	C-term Strep II
3	None	pRS-424	No tags

Table 3.2 **Ndc10 plasmids.** Plasmids constructs used to express and purify different forms of Ndc10.

(GE Healthcare) and eluted in the gel filtration buffer composed of 15 mM Tris-Cl pH 8, 200 mM NaCl, 2 mM DTT and 10 μM $ZnCl_2$.

The SDS-PAGE after IMAC (Fig 3.6) showed three bands present at the expected molecular weights for Cep3, Ctf13 and Skp1. The band for Cep3 was more intense than the rest of the subunits because the stoichiometry of the complex means there twice as much Cep3 as there is of Ctf13 and Skp1 as well as Cep3 expressing better than the other subunits. After Heparin chromatography, the SDS-PAGE and the chromatogram shows two populations emerging, excess Cep3 and the CBF3 core complex. The gel filtration leads to two separate species with the initial fractions containing the CBF3 core complex with some less populated contaminants and the later fractions containing the core complex but with an additional intense contaminants emerging. The intensity of the contaminant is inverse to the intensity of the Cep3 band suggests; the contaminants could be a degradation product of Cep3.

3.5 Protein Purification of the Ndc10

Protocols had been established to express and purify the Ndc10 DNA binding domain and the Ndc10 DNA binding domain with the dimerisation domain in *E. coli* systems (Cho and Harrison, 2012, Perriches and Singleton, 2012). Since *E. coli* yields are generally low for Ndc10 constructs and since endogenous protein modifications may be relevant to Ndc10 function (Montpetit et al., 2006, Peng et al., 2011), one aim of the project was to develop a purification for Ndc10 using the *S. cerevisiae* expression system (Fig 3.7 and 3.8). The domain boundries were determined *in silico* by Jonathan Lucas who also cloned the constructs for protein expression.

Ndc10 was expressed in *S. cerevisiae* (Section 2.5.2) and lysed (Section 2.6.2). After the cells were lysed and pH adjusted, the lysate was sonicated to shear the excess DNA, for 3 cycles at an amplitude of 65 % for a total time of 2 minutes, with pulses on for 2 seconds and off for 4 seconds. The final cycle was carried out using the same settings at an amplitude of 80 %. After clearing the lysate, the sample was sonicated again for 3 cycles at an amplitude of 30 % for a total time of 1 minute, with pulses on for 2 seconds and off for 4 seconds. This enabled the loading of the cleared lysate on to a HiTrap^{TM} column without blocking it.

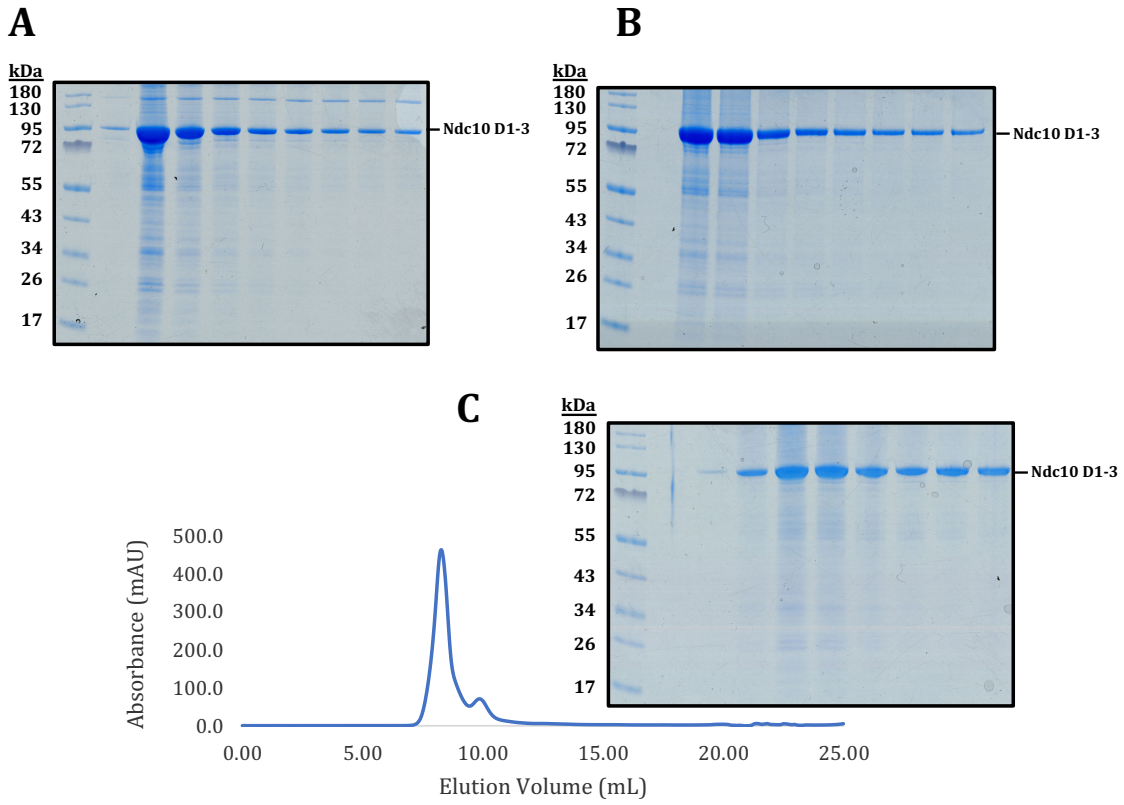


Fig. 3.7 **Protein purification of Ndc10 domains 1-3.** (A) SDS-PAGE of the elution after 5 mL Strep affinity chromatography. (B) SDS-PAGE after Heparin chromatography. (C) SDS-PAGE and chromatogram from gel filtration of Ndc10 domains 1-3.

The sample was first loaded on to a 5 mL STREPTrap^{TM} HP with a Peristaltic Pump P-1 (GE Healthcare) which was pre-equilibrated with lysis buffer (25 % Glycerol, 40 mM Tris-Cl pH 8.0, 1 M NaCl, 2 mM Magneisium acetate, 1 mM EDTA pH 8, 0.2 % IGEPAL CA-630 (Sigma), 40 mM NaF, 4 μg/mL DNase I, 3 mM DTT

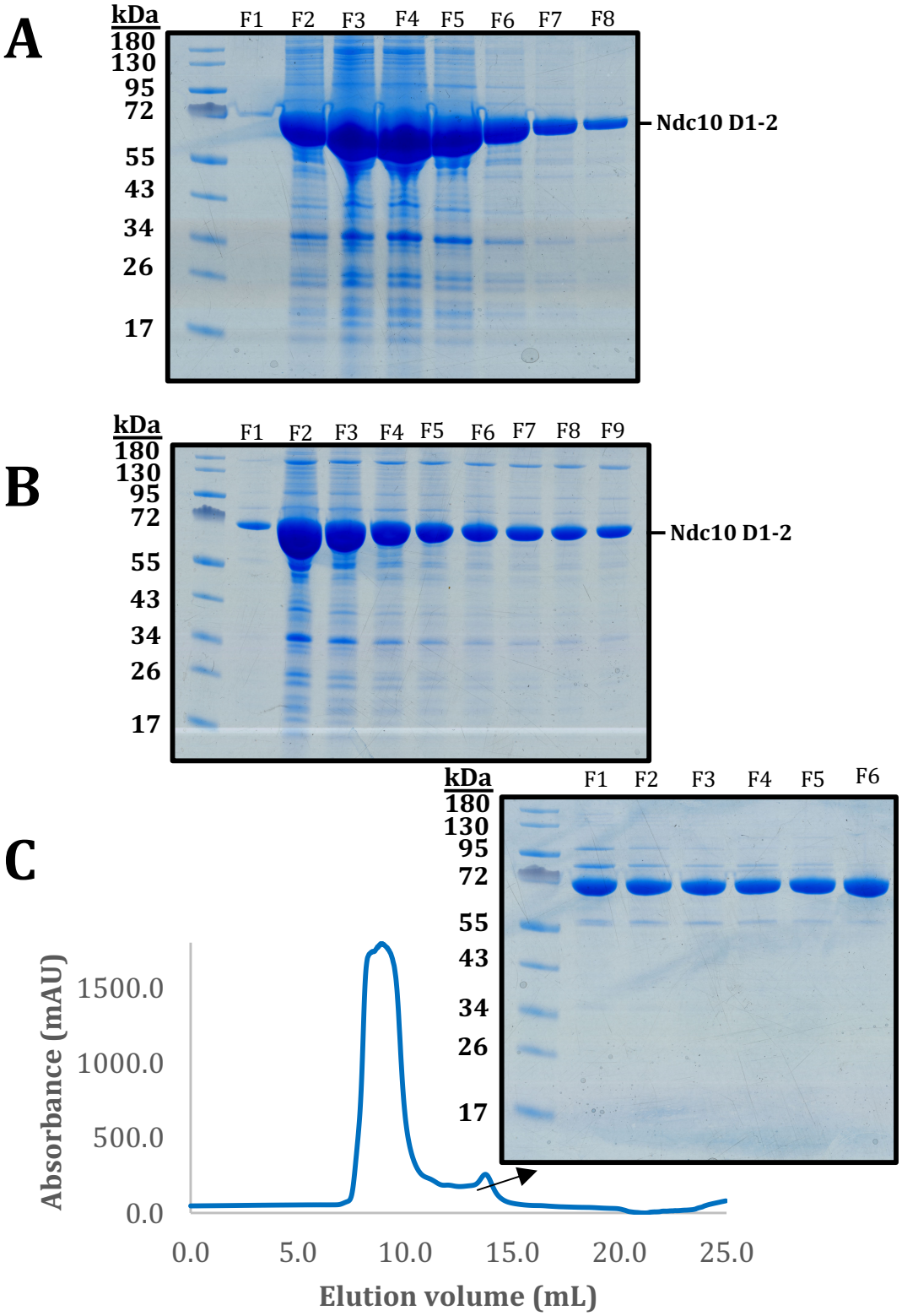


Fig. 3.8 **Protein purification of Ndc10 domains 1-2.** (A) SDS-PAGE of the elution after 5 mL Strep affinity chromatography. (B) SDS-PAGE after Heparin chromatography. (C) SDS-PAGE and chromatogram after gel filtration of Ndc10 domains 1 to 2.

and 2 cOmplete^{TM}, Mini, EDTA-free Protease Inhibitor Cocktail tablets (Roche)). The protein bound to the column was washed with Buffer A (20 mM Tris-Cl pH 8, 500 mM NaCl, 2 mM DTT, 1 mM EDTA) and eluted in Buffer B (20 mM Tris-Cl pH 8, 500 mM NaCl, 2 mM DTT, 1 mM EDTA, 3 mM Desthiobiotin). After the STREP affinity purification, there is a yeast contaminants, which co-elutes at high concentrations. This contaminant co-elutes with the Ndc10 and cannot be separated using ion exchange and gel filtration, but can be separated from the protein using Heparin chromatography. Therefore after STREP affinity purification the protein sample has the NaCl concentration adjusted to 200 mM and loaded on to a 1 mL HiTrap® Heparin HP (GE Healthcare) using a Peristaltic Pump P-1 (GE Healthcare), which was pre-equilibrated with 5 column volumes of Buffer HA (20 mM Tris pH 8, 10 % glycerol, 2 mM DTT), 10 column volumes of Buffer HB (20 mM Tris pH 8, 10 % glycerol, 2 mM DTT and 1M NaCl), followed by 15 column volumes of 20 % Buffer HB, which enables the salt concentration in the column to match the salt concentration of the sample. The resin containing the bound protein was washed with 20 % Buffer HB, to remove contaminants for 10 column volumes. The sample was eluted in Buffer HB. The protein sample was concentrated using Pierce^{TM}TM Protein Concentrators PES (Thermo Scientific^{Tm}). The PES membrane had lower non-specific interactions with Ndc10 than the regenerated cellulose membranes of Amicon® Ultra, which resulted in loss of Ndc10. The sample was loaded on to a Superdex Increase 200 10/300 GL (GE Healthcare) and eluted in the gel filtration buffer composed of 15 mM Tris-Cl pH 8, 200 mM NaCl, 2 mM DTT. This protocol was used to purify Ndc10 domains 1-3 (Fig 3.7) and Ndc10 domains 1 to 2 (Fig 3.8).

3.6 Protein Purification of the Truncated CBF3 complex

Different research groups have faced issues purifying the CBF3 complex where Ndc10 remains associated with the CBF3 core complex (Leber et al., 2018, Zhang et al., 2018). Previously, Ndc10 domains 1-2 association with the core CBF3 complex was studied by reconstituting the two components *in vitro*. It is important to develop a method to co-express and purify the complex to obtain a more physiologically relevant protein complex.

Construct	**Proteins**	**Vector and selection marker**	**Protein tags**
1	Cep3ΔN (Residues 47-608) Skp1 Ndc10 domains 1-2 (Residues 1-554)	pRS-426 TRP	Cep3 (N-term 6x His) Ndc10 (C-term Strep II)
2	Cep3ΔN (Residues 47-608) Skp1 Ndc10 domains 1-3 (Residues 1-701)	pRS-426 TRP	Cep3 (N-term 6x His) Ndc10 (C-term Strep II)
3	Sgt1 Ctf13	pRS-424 URA	Sgt1 (C-term Strep II) Ctf13 (C-term CBP)

Table 3.3 **Truncated CBF3 complex expression plasmids.** Plasmids constructs used to express and purify truncated CBF3 complexes. These plasmids were gifted By Dr. Wenjuan Zhang and Jonathan Lucas.

A major issue in preventing the Ndc10 from separating from the core complex is that Ndc10 has higher affinity for specific resins than it does for the CBF3 core complex subunits. By testing different anion exchange chromatography columns such as MonoQ, Q-sepharose and HiTrap Q, only ResourceQ^{TM} column was able to keep the complex intact. The ResourceQ^{TM} column was the only ion exchange column to be tested which did not have sepharose based resin, suggesting that Ndc10 has strong non-specific interactions with sepharose.

A purification strategy was devised to avoid the use of resins which Ndc10 would have strong affinity for. This includes avoiding the use of the C-terminal Strep II tag at the C-terminus of Ndc10 domains 1-3. Due to time constraints Calmodulin affinity purification was used, where the calmodulin is bound to a sepharose based resin, as this is a very well established protocol in the lab.

To purify the co-expressed truncated CBF3 complex, the lysate was thawed, re-suspended in 2x Lysis buffer (25 % Glycerol, 50 mM Tris-Cl pH 8.0, 500 mM NaCl, 2 mM Magneisium acetate, 2 mM Imidazole, 4 mM $CaCl_2$, 0.2 % IGEPAL CA-630 (Sigma), 40 mM NaF, 4 μg/mL DNase I, 4 mM DTT and 2 Protease Inhibitor Cocktail tablets (Roche)). After adjusting pH to 8, the cell lysate was then sonicated for a total

of 2 minutes at 65% amplitude, with pulses on for 2 seconds and off for 4 seconds for three cycles. The fourth cycle was conducted in the same manner except the amplitude was increased to 85%.The lysate was cleared using ultracentrifugation at 38,000 rpm at 4°C for 1 hour 30 minutes.The cleared lysate was incubated with 900 μL of Calmodulin – Sepharose resin for 12-16 hours at 4°C, on a roller. The resin was washed with Buffer A (10 mM Tris-Cl pH 8, 500 mM NaCl, 1 mM Magnesium Acetate, 1 mM Imidazole, 2 mM Calcium chloride, 0.1 mM EGTA pH 8, 2 mM DTT) and eluted in Buffer B (10 mM Tris-Cl pH 8, 150 mM NaCl, 1 mM Magnesium Acetate, 1 mM Imidazole, 2 mM Calcium chloride, 10 mM β-mercaptoethanol, 6 mM EGTA). The eluted protein was loaded on to 1 mL Resource^{TM} Q (GE Healthcare) using a 50 mL Superloop (GE Healthcare). The column was equilibrated with 5 column volumes of Q Buffer A (20 mM Tris-Cl pH 8, 10% Glycerol, 2 mM DTT), 10 column volumes of Q Buffer B (20 mM Tris-Cl pH 8, 1M NaCl, 10 % Glycerol, 2 mM DTT) and then 10 column volumes of 15 % Buffer B. The protein was eluted across a linear gradient of 15 % to 50 % Buffer B. The protein was concentrated using Vivaspin 20 centrifugal concentrator (Sartorius). The protein was dialysed into Buffer composed of 20 mM Tris- Cl, 200 mM NaCl, 1 mM $MgCl_2$ and 2 mM DTT using a Slide-A-Lyzer MINI Dialysis Device (ThermoFisher) for 12 hours at 4°C. The sample was then incubated with annealed 56 base pair CDEIII DNA, for 45 minutes at room temperature. The sample was applied to Superdex® 200 10/300 GL (GE Healthcare) and eluted in buffer containing 15 mM Tris-Cl pH 8, 200 mM NaCl, 1 mM $MgCl_2$, 2 mM DTT. The protocol and buffers were used to purify both truncated CBF3 complex with Ndc10 domains 1 to 2 (Fig 3.9) (Constructs 1 and 3) (Table 3.3) and Ndc10 domains 1-3 (Fig 3.10) (Constructs 2 and 3) (Table 3.3). To purify the complex without DNA, the dialysis step was omitted and the sample was gel filtrated.

3.7 Discussion

CBF3 Core complex

A strategy was devised to purify a sufficient yield of the CBF3 core complex. The proposed method overcomes a number of observed challenges which were not present when purifying the CBF3ΔN (Section 2.6.11). These challenges include imidazole sensitivity, salt concentration sensitivity, Ctf13 affinity tagging and instability at high protein complex concentrations. The imidazole sensitivity is consistent with observations made when purifying Cep3 expressed in *E. coli*, and an additional study

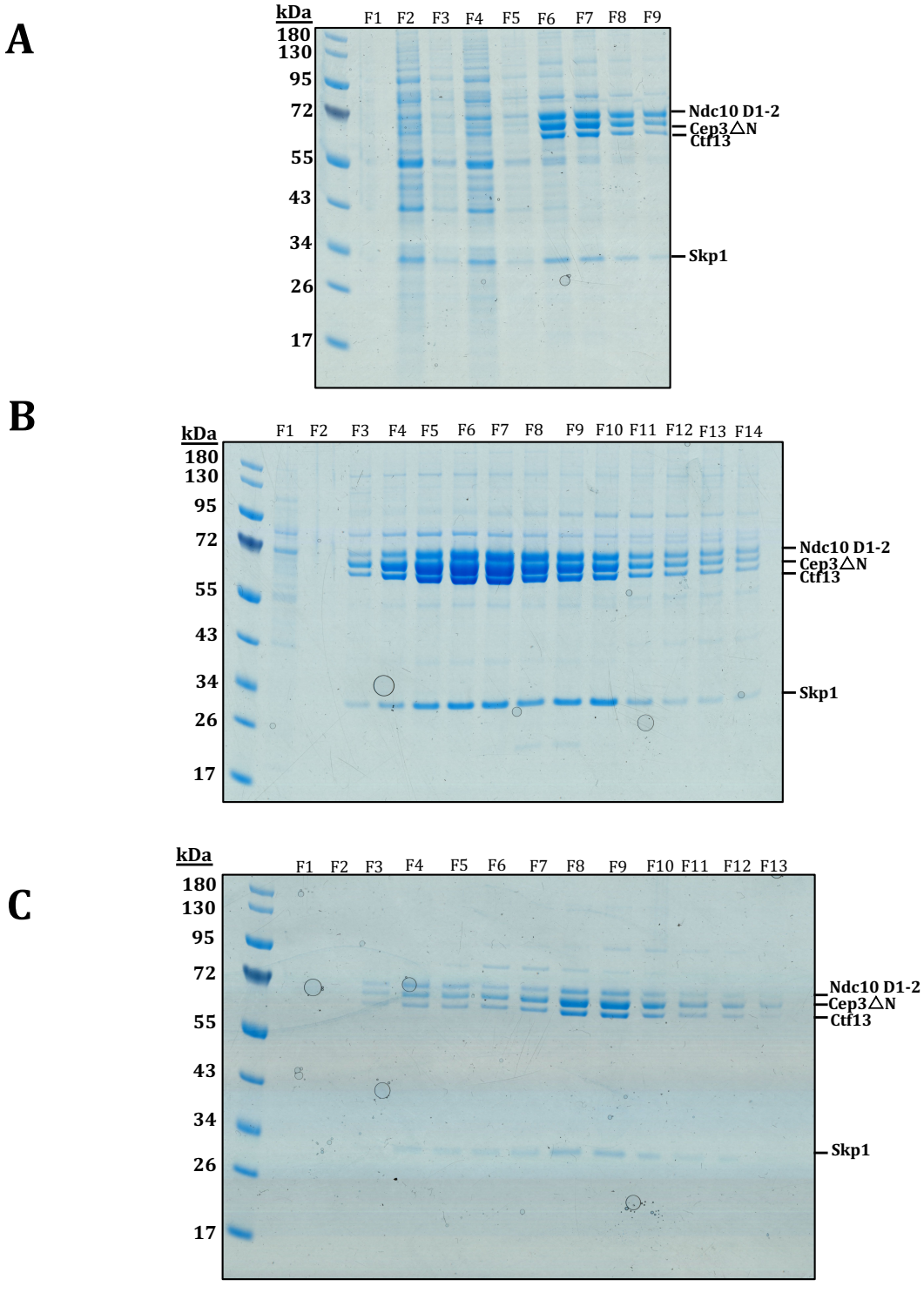


Fig. 3.9 **Protein purification of the truncated CBF3 complex with Ndc10 domains 1 to 2.** (A) SDS-PAGE of the elution after Calmodulin affinity purification. (B) SDS-PAGE after anion exchange chromatography. (C) SDS-PAGE after gel filtration Ndc10 domains 1 to 2.

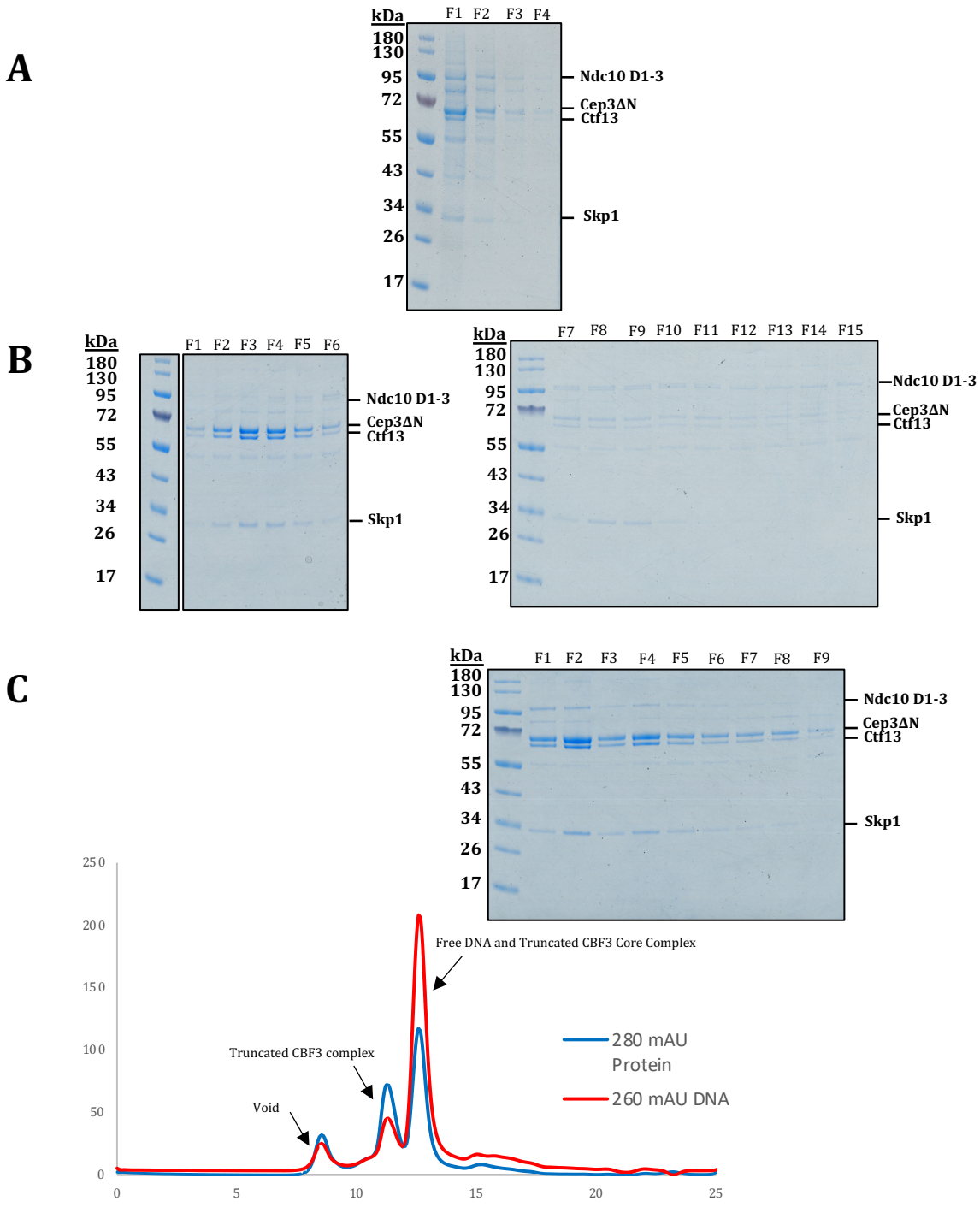


Fig. 3.10 **Protein purification of the truncated CBF3 complex with Ndc10 domains 1-3.** (A) SDS-PAGE of the elution from Calmodulin affinity purification. (B) SDS-PAGE after anion exchange chromatography. (C) SDS-PAGE and the chromatogram with 260 nm absorbance (red) and 280 nm absorbance (blue) after gel filtration.

eluted Cep3 from Nickel-affinity resin using EDTA (Purvis and Singleton, 2008). This strategy avoids the use of EDTA to ensure the Zn_2Cys_6 cluster domain is not affected by the chelating agent (NYBORG and PEERSEN, 2004). Anion exchange chromatography was not used as the CBF3 core complex was salt sensitive in our hands, however other groups have used it successfully although the salt concentration used in that study was not stated (Leber et al., 2018)

The purification procedure shows that the CBF3 core complex is tightly associated. A number of the challenges associated with the purifcation provide an insight into the underlying biology of the CBF3 complex and how it functions. It is evident that observed challenges result from the introduction of the Zn_2Cys_6 cluster domain, as when this domain is truncated, these problems are no longer present (Section 2..6.11). The increased salt sensitivity could be an indication of the role the Zn_2Cys_6 cluster domain plays in DNA binding. DNA binding domains typically have large proportion of positively charged residues to drive electrostatic interactions which play a role in DNA recognition and binding (Jones et al., 2003, Temiz et al., 2010). Approximately 26% of the total residues of the Cep3 Zn_2Cys_6 cluster domain are either lysine and arginine, where that figure is only 12% for the rest of Cep3, Ctf13 and Skp1. The stability issue could be a result of the need for higher salt concentrations due to the increased number of charged residues contributed by the Zn_2Cys_6 cluster domain (Lindman et al., 2006). The stability issues at higher concentrations of CBF3 core complex was exemplified in an experiment conducted (not shown), where the protein complex was expressed in 48 L of cell culture and the protein eluted from the Nickel-affinity resin as precipitate. Previous studies have suggested that phosphorylation and ubiquitination pathways work in synergy to regulate cellular levels of CBF3 (Kaplan et al., 1997). The heterodimer Ctf13-Skp1, present in the core complex, are speculated to play a role in this regulation (Kaplan et al., 1997). This was said to ensure the CBF3 complex assembly and disassembly was coupled to the replication of centromeric DNA in S phase and likely prevents ectopic kinetochore assembly (Kaplan et al., 1997). The instability observed at higher concentration could reflect the need for regulation of CBF3 levels, and the precipitation relates to a secondary mechanism to control CBF3 concentration.

Ndc10

Previous expression and purification of Ndc10 proteins were conducted using *E. coli* for structural studies (Cho and Harrison, 2012, Perriches and Singleton, 2012). This project presents a methodology to express and purify the protein in *S. cerevisiae*, which

will enable future biochemical experiments to be conducted on the protein which are more physiologically relevant containing the correct post-translational modifications. Optimisation of the process showed expression of Ndc10 requires repression of the promoter to inhibit leaky expression, to ensure the cells grow to the optimum density. Also, Ndc10 domains 1-2 showed greater stability than Ndc10 domains 1-3, with Ndc10 domains 1-2 yield significantly higher than Ndc10 domains 1-3. It should be noted that the secondary structure prediction data suggests Ndc10 to be structured whilst domains 3-5 contain a large amount of intrinsically disorderd sequences (Perriches and Singleton, 2012). In experimental terms, this could have occurred as a result of either Ndc10 domains 1-3 having a greater degree of flexibility than domains 1-2, resulting in protein loss to the membrane of the protein concentrator during the purification or as a result inaccurate domain boundaries.

The observations made during the purification protocol provides speculative insight into likely Ndc10 functions. The leaky expression affecting the growth of the cells was observed when Ndc10 is expressed both alone or when expressed with the other subunits. The instability that Ndc10 imparts relates to domains 1-3 as domains 4 and 5 were not present. One possibility is that the Ndc10 either associates with the other co-expressed subunits or the endogenous CBF3 subunits and drives centromeric DNA association. However, as both domains 4 and 5 of Ndc10 are not present, this prevents the recruitment of Cse4 nucleosome and could inhibit recruitment of subsequent kinetochore components. Alternatively, the presence of higher levels of Ndc10 could affect other processes associated with Ndc10 including anaphase spindle stability and cytokinesis (Bouck and Bloom, 2005). Future experiments where full-length Ndc10 is expressed using the same protocol, could be used as a comparison, as domains 4 and 5 would be present. These ideas are speculative and further experiments would need to be conducted to isolate the cause of Ndc10 leaky expression affecting cellular growth. Another possibility for the observed growth defect could be related to the Ndc10 DNA binding function. Ndc10 is known to bind to DNA in a non-specific manner and thus the expressed Ndc10 could be non-specifically binding to different DNA segments and interfering with lots of different cellular processes (Espelin et al., 2003, Cho and Harrison, 2012). Moreover, the observed stability difference between Ndc10 domains 1-2 and Ndc10 domains 1-3 could be related to how Ndc10 functions. It can be speculated that Ndc10 domains 4-5 play an additional role in stabilising the dimerisation domain or contribute to the dimerisation activity. The establishment of this protocol enables future experiments to be conducted on both these Ndc10 truncations and full-length Ndc10, to determine the biological reasoning as to why

Ndc10 leaky expression affects cellular growth and why the Ndc10 truncations show different levels of stability.

Truncated CBF3 complex

Previous attempts at purifying an intact CBF3 complex resulted in the complex separating into core CBF3 complex and Ndc10 (Leber et al., 2018, Zhang et al., 2018). The purification strategy devised was able to yield an intact truncated CBF3 complex containing either Ndc10 domains 1-2 or Ndc10 domains 1-3. Both truncated CBF3 complexes showed that the Ndc10 readily disassociated from the core complex over the course of the purification method, but this was more striking for Ndc10 domains 1-3. This result was consistent with the observations made in the CBF3-CEN3 study where the complex was supplemented with additional Ndc10 to compensate for the disassociation. (Yan et al., 2018).

The disassociation of Ndc10 from the CBF3 core complex can be explained by the relatively small interaction surface between Ndc10 and Ctf13 (Zhang et al., 2018, Yan et al., 2018). The weak interactions explains why previous cryo-EM studies on the CBF3 complex only yielded CBF3 core complex and not the full complex (Leber et al., 2018). The ability for Ndc10 to readily disassociate from the CBF3 core complex is of interest because the rest of the CBF3 complex remains tightly associated. This raises new questions about the ability of Ndc10 to readily disassociate from the the rest of the complex and how this may relate to the CBF3 assembly pathway and function. Possibilities for the weak interactions between CBF3 core complex and Ndc10 include, Ndc10 playing a transient role in CBF3 function, where Ndc10 disassociates after Cse4 nucleosome recruitment or that centromeric DNA is needed to stabilise the association of Ndc10 with the rest of the complex. The established methodology in this study will enable physiologically relevant CBF3 biochemical studies to probe these new questions further.

Chapter 4

Results 2: Structural analysis of the truncated CBF3 complex

4.1 Introduction

Prior to the project, structural insight into the CBF3 complex was limited to the individual subunits of Cep3, Skp1 and Ndc10 (Schulman et al., 2000, Bellizzi et al., 2007, Purvis and Singleton, 2008, Cho and Harrison, 2012, Perriches and Singleton, 2012, Willhoft et al., 2017). Recently, a number of cryo-EM structures relating to the CBF3 core complex and the CBF3 core complex with Ndc10 DNA binding domain, were released (Leber et al., 2018, Zhang et al., 2018). These publications characterised the structure of Ctf13 for the first time and provided details related to subunit interactions (Zhang et al., 2018). However, the published structures lacked DNA bound to the complex (Leber et al., 2018, Zhang et al., 2018). The DNA binding function of the complex is inferred through a combination of modelling linear DNA and by analysing the electrostatic potential of the channel and the composition of the residues protruding into the channel (Zhang et al., 2018). Information related to how the complex binds DNA, the conformation of the protein complex on the DNA and how the presence of DNA affects the position of the subunits relative to one another is required. In addition, structural detail of Ndc10 was limited to the DNA binding region of Ndc10 (Domains 1-2) with domains 3 to 5 missing (Fig 1.15) (Cho and Harrison, 2012, Perriches and Singleton, 2012, Zhang et al., 2018). As domain 3 facilitates dimerisation of Ndc10,

it is not known how the second Ndc10 protomer would position itself relative to the other subunits and the DNA (Cho and Harrison, 2012, Perriches and Singleton, 2012).

This chapter describes single particle cryo-EM analysis conducted on the truncated CBF3 complex with 56 bp CDEIII DNA. The truncated CBF3 complex is composed of the CBF3ΔN core complex where Cep3 has residues 1-46 truncated and Ndc10 domains 1-3, with domains 4 and 5 truncated. The N-terminal region of Cep3 was removed to improve stability of the protein complex (see Section 2.6.11). Ndc10 domains 1-3 was used because these domains were sufficient for structural understanding of DNA binding and oligomerisation and co-expression with the full-length protein failed.

4.2 Sample preparation

The truncated CBF3 complex was co-expressed as described in Section 2.5.3, purified and bound to DNA as detailed in Section 3.6 (Fig 3.10). The chromatogram generated from the gel filtration (Fig 3.10.C) shows multiple distinct peaks, one of which was the void peak. The protein complex peak was identified by overlaying this chromatogram with gel filtration chromatograms for the CBF3ΔN core complex, 56 bp CDEIII DNA, Ndc10 domains 1-3 and the truncated CBF3 complex not bound to DNA (Fig 4.1)

The overlay of the chromatograms helped define each peak and identified the peak at 11.5 mL (Fig 4.1.A) as the peak containing the sample. This sample was then applied to cryo-EM grids for vitrification and cryo-EM data collection (Fig 2.10). Further analysis of the absorbance at 260 nm (Fig 4.1.B), used to determine the quantity of DNA, indicates that the DNA has either eluted without binding to the protein complex or has bound to the CBF3ΔN core complex. The absorbance at 260 nm, at the position of the protein complex elution, was relatively low suggesting very little DNA was present amongst the protein complexes. The void peak at 8 mL, shows the absorbance at 260 nm and 280 nm are equal at this position, which suggests some of the DNA bound complex may have aggregated. Negative stain of the void volume showed a mix of particles of various different sizes, meaning it was not suitable for structural analysis (data not shown). The micrographs collected (Fig 4.2) showed sufficient quantity of particles but these particles were clustered together.

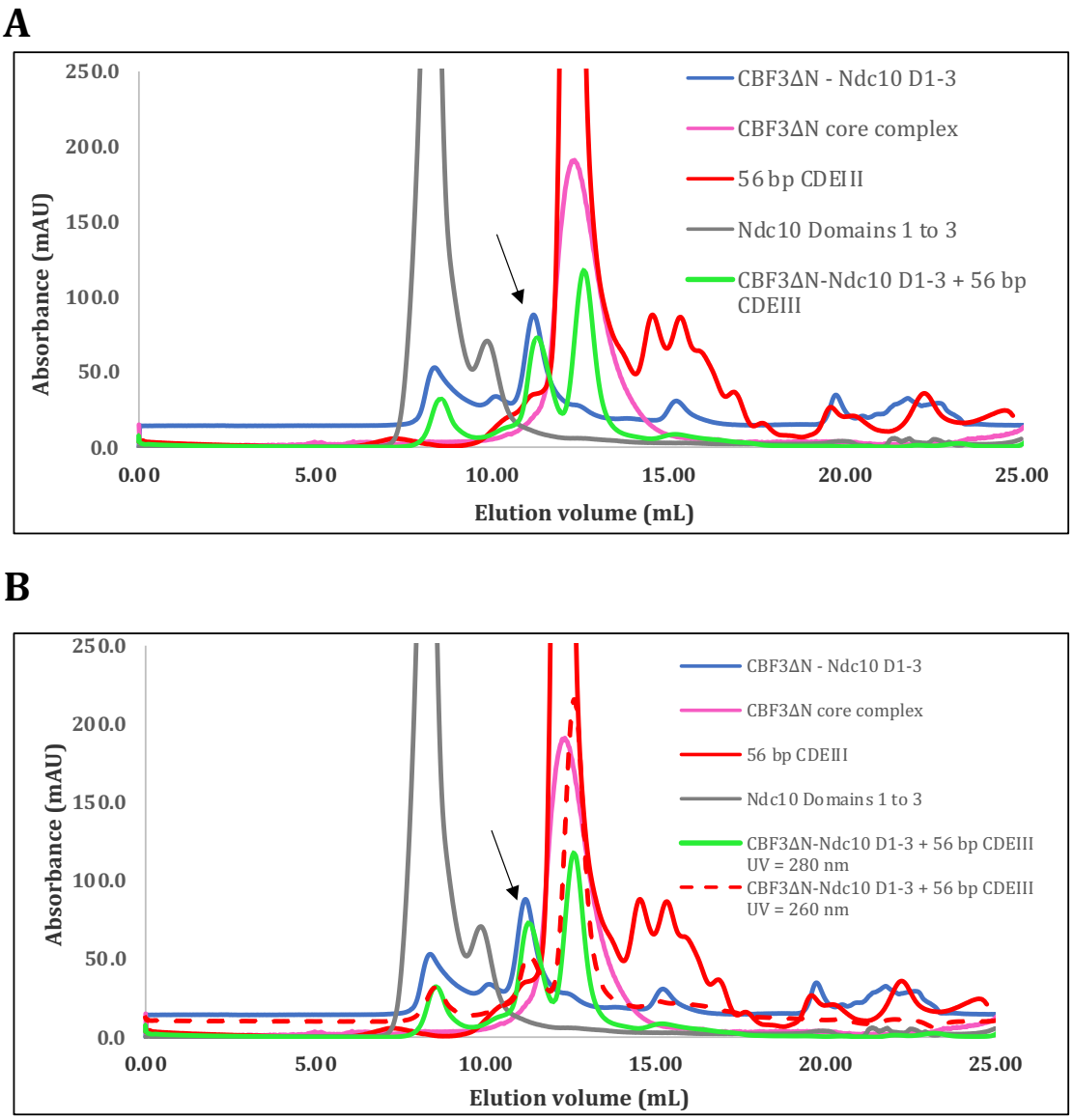


Fig. 4.1 **Combined chromatograms to identify the position of the truncated CBF3 complex - 56 bp CDEIII.** (A) Overlay of the gel filtration chromatograms, at UV absorbance of 280 nm. Arrow highlights the the area of the peak used for structural analysis.(B) Overlay of the gel filtration chromatograms from (A) with the addition of the chromatogram from UV aborbance of 260 nm, from the truncated CBF3 complex with DNA. The arrow highlights the the area of the peak used for structural analysis.

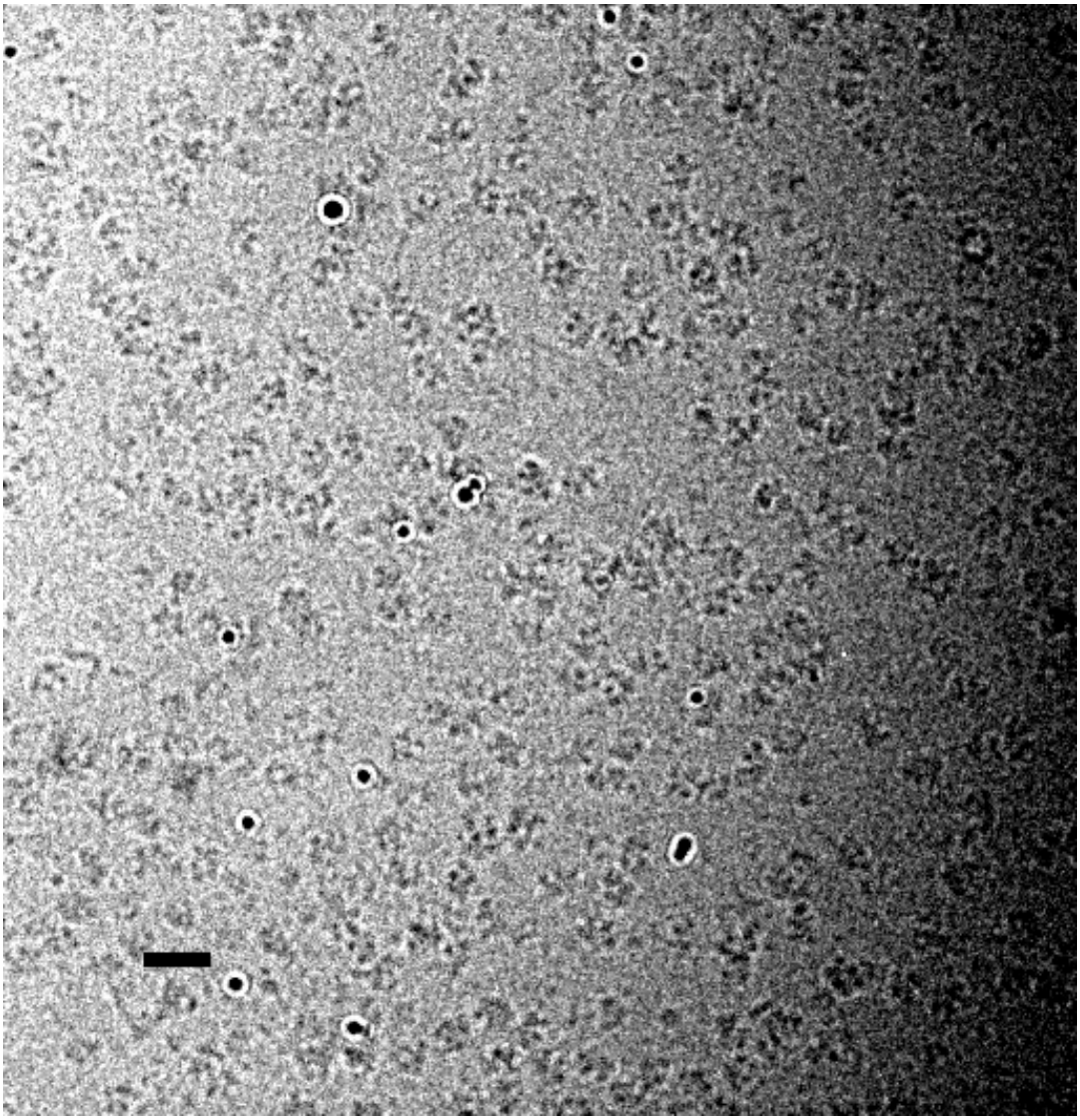


Fig. 4.2 **Micrograph from the cryo-EM data collection.** An example of a micrograph of the truncated CBF3 complex from the collected data set. The scale bar represents 20 nm.

4.3 Dimeric CBF3 complex

During the process of cryo-EM data collection, a structure of the CBF3 complex bound to CEN3 DNA was published, with an unexpected conformation (Yan et al., 2018). The insect cell expressed structure showed that in the presence of Ndc10 and the 147 bp CEN3 DNA, the CBF3 complex is composed of two core complexes and two copies of Ndc10 (Yan et al., 2018). The CBF3 complex occupies 6 bp of CDEII, the entire

CDEIII and an additional 24 bp to the right of CDEIII (Yan et al., 2018). The 56 bp occupied by the CBF3 complex matches the 56 bp CDEIII determined using DNase I cleavage mapping and is the exact sequence used in this experiment (Lechner and Carbon, 1991). The conformation was unexpected because the biochemistry and genetic studies conducted on the CBF3 complex to calculate the stoichiometry suggested only one CBF3 core complex present with two copies of Ndc10 (Lechner and Carbon, 1991, Espelin et al., 1997). This Section will investigate whether this unexpected CBF3 conformation is present in this data set. The search for the dimeric CBF3 complex used two different approaches; a template based particle picking approach and a non-template based particle picking approach.

4.3.1 Non-template picking and processing

After completing the motion correction and CTF correction, particle picking was performed using Gautomatch v0.56. The picking was performed without any templates, with a particle diameter of 340 Å specified for the particle identification. This value was calculated in Chimera using the CBF3-CEN3 cryo-EM map (EMDB:0096).Subsequent processing steps were performed using Relion-3.0.

Gautomatch selected 39,154 particles and these particles were extracted using a box size of 400 Å. Two rounds of 2D classification was then performed to remove noise as described in Section 2.10.2. The 2D classification for this experiment was performed template-free and using a mask diameter of 360 Å. The first round of 2D classification was performed with 200 classes and the second round was performed with 100 classes, due to the reduced number of particles.

After two rounds of 2D classification, 22 classes were selected, consisting of 16,912 particles (Fig 4.3). The 2D classes selected for 3D processing shows two particles in close proximity. The density was not connected, unlike 2D classes of the dimeric CBF3 complex (Yan et al., 2018). In addition the channel of the two core CBF3 complexes are not facing each other in some of the classes, unlike what would be expected for the bilobal structure. Initial analysis of the 2D-classes suggests that, because the connecting density is not present there is a possibility that particle picking software has boxed two particles in close proximity, as a result of the particles congregating next to one another in the micrograph (Fig 4.2). However, the connecting density may also not be visible because the Ndc10, which would be expected to connect the two

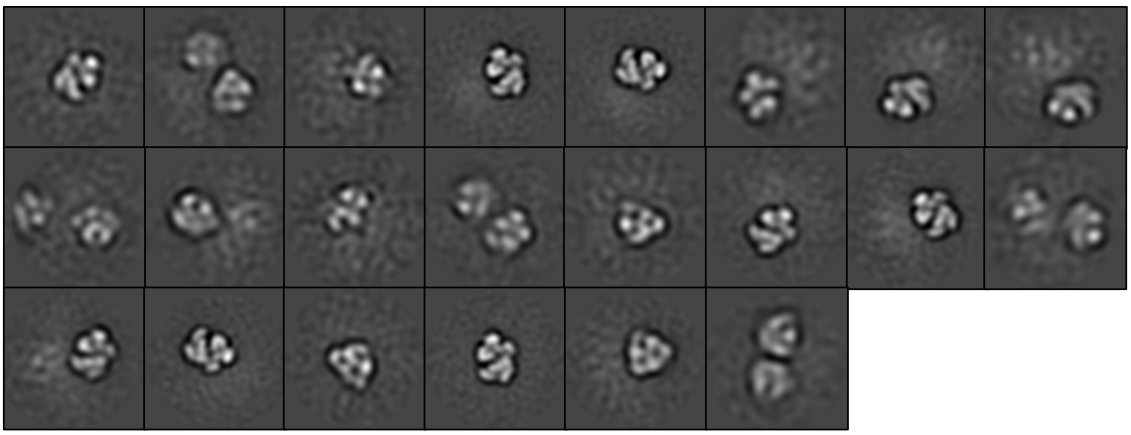


Fig. 4.3 **2D classification of particles selected without template.** The 2D classes of particles selected without using templates for particle picking, selected to perform the subsequent 3D image processing.

core CBF3 complexes, may be very flexible, resulting in the orientation of the core complexes being different in each particle, leading to the software being unable to align the particles.

3D classification was subsequently performed on these selected particles from the 2D analysis (Fig 4.3), as described in Section 2.10.2. The number of classes specified for 3D classification was lowered to 3, due to the low number of particles available for processing. Particle diameter was specified as 360 Å. Two separate 3D classifications was performed using two different initial models. The first initial model was generated using Relion-3.0 SGD algorithm, and this model was subsequently low pass filtered to 60 Å and used as a reference map (Fig 4.4). The second initial model was the atomic model of CBF3-CEN3 PDB structure (PDB:6GYS). Using Eman pdb2mrc function, a reference structure was generated, with the pixel size adjusted to 0.85 Å and the box size adjusted to 400 Å. This structure was then low pass filtered to 60 Å and used as a reference map for 3D classification.

3D classification with SGD generated model

The low resolution of the 3D classes, meant map analysis was challenging (Fig 4.4). Analysis of the three 3D volumes shows Class 1 has connecting density between two larger volumes. The positioning of the densities suggests the two large volumes could be the CBF3 core complex and the connecting density is likely to be Ndc10 domains 1-3. To confirm if the density assignment was correct, the 3D-map for Class 1 was

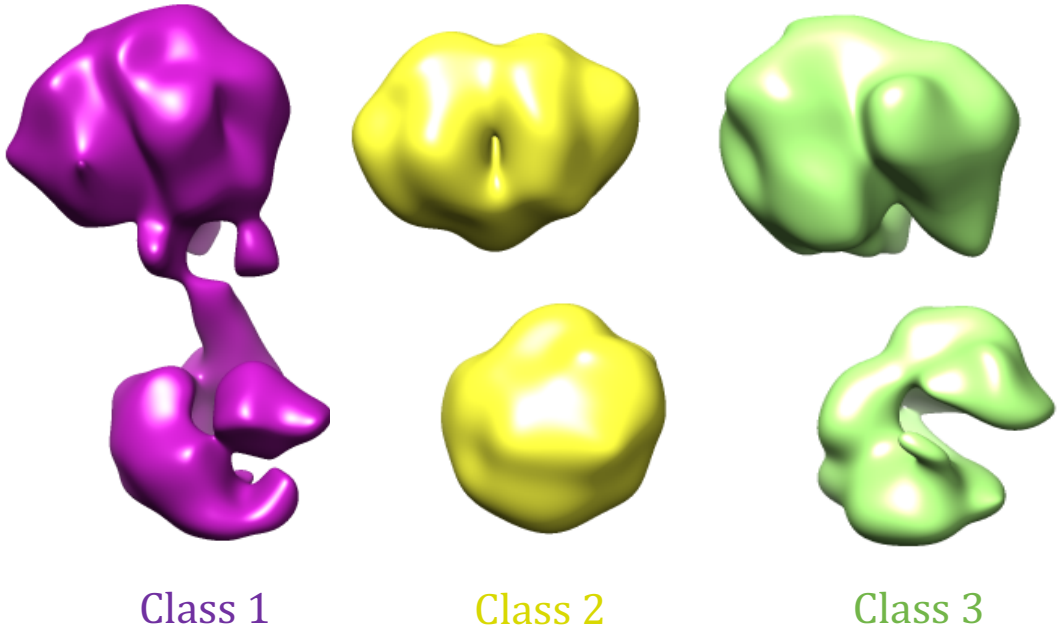


Fig. 4.4 **3D classification of particles picked without template, conducted with SGD generated initial model.** Three 3D volumes generated from the 3D classification performed using the SGD generated intiial model as a reference map.

Class	Class distribution %
1	42.3
2	32.6
3	25.1

Table 4.1 **Class distribution of the 3D classes performed with SGD generated initial model on the non-template based picked particles.** The table lists percentage of the total particle count occupying each class.

manually fitted and aligned to CBF3-CEN3 map using Chimera (Fig 4.5). Fitting the two 3D maps, showed the distance between the two large densities in Class 1 was identical the distance between the CBF3 core complexes in the CBF3-CEN3 map (Fig 4.5). The size of Ndc10 connecting density was also identical, but the orientation of the Ndc10 density differed between the two maps. Class 1 is also the most populated class, with 7153 particles (Table 4.1) and combined with the analysis conducted above, this class was selected for 3D-refinement.

3D refinement of Class 1

3D-refinement was performed on Class 1 (Fig 4.4) as described in Section 2.10.2. The

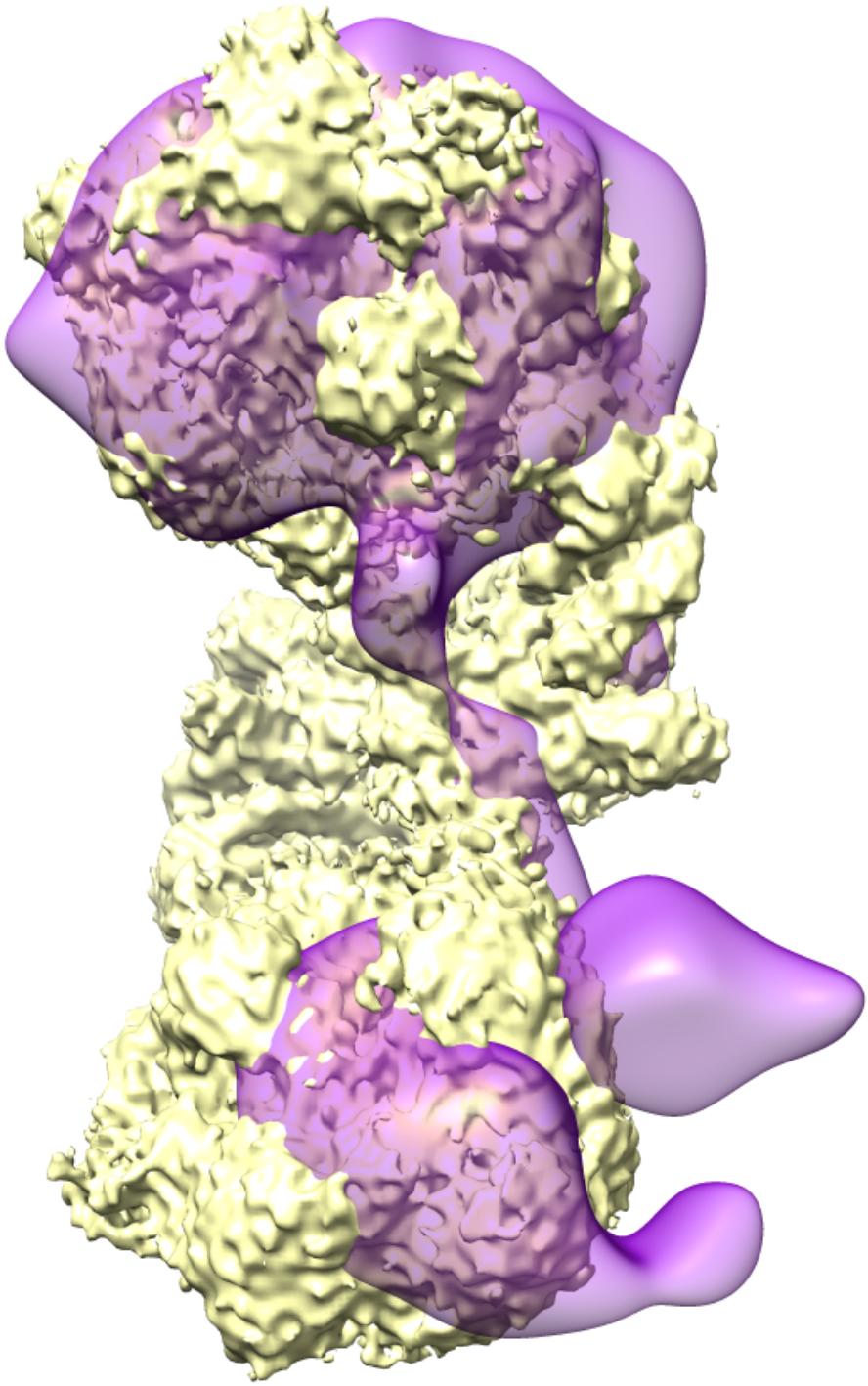


Fig. 4.5 **Map comparison - Class 1 and CBF3-CEN3.** Class 1 from the 3D-classification (purple) performed on particles picked without template, using the SGD initial model, was aligned and fitted into the CBF3-CEN3 map (yellow) using the Chimera software. EMDB:0096

7243 particles from class 1 were selected and a mask diameter of 360 Å was specified. The 3D map of Class 1 was used as a reference map for the refinement. The refined 3D map showed that the map quality failed to improve (Fig 4.6). The appearance of the map as well as the angular distribution showed preferential orientation of particles. This likely caused by the small number of particles provided for refinement, and a larger number of particles may be required to obtain additional particle views.

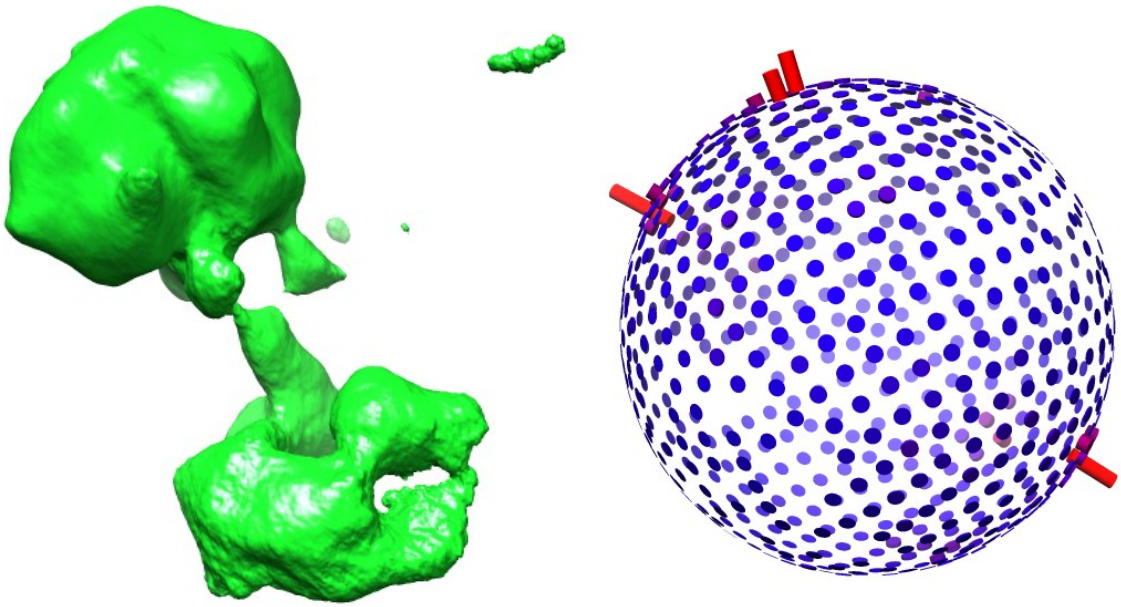


Fig. 4.6 **3D auto-refinement of Class 1.** The refined 3D map of class 1 and the angular distribution of the particles used to generate the 3D-volume.

3D-classification with published CBF3 - CEN3 as reference map

The second 3D-classification of the non-template picked particles (Fig 4.3) was performed using the low-pass filtered CBF3-CEN3 atomic model (PDB:6GYS) as described above. Two of the three 3D-classes were similar in appearance (Fig 4.7). The two classes showed the core CBF3 complex with a tail next to additional density not connected. The second density is likely to be another core CBF3 complex. The two classes were less populated than Class 3 (Table 4.2). To assign the densities, the Class 1 and Class 2 maps (Fig 4.7) were manually aligned and fitted into CBF3-CEN3 map (Fig 4.8). The alignment and fitting of the maps showed the distance between the two core CBF3 complexes was identical to the CBF3-CEN3 map. The connecting density in Class 1 aligned with the Ndc10 of the CBF3-CEN3 map, suggesting the Ndc10

orientation is identical in the two 3D-maps (Fig 4.8). The connecting density in Class 2 was of the same size as the Ndc10 of CBF3-CEN3, but the orientation was different (Fig 4.8)

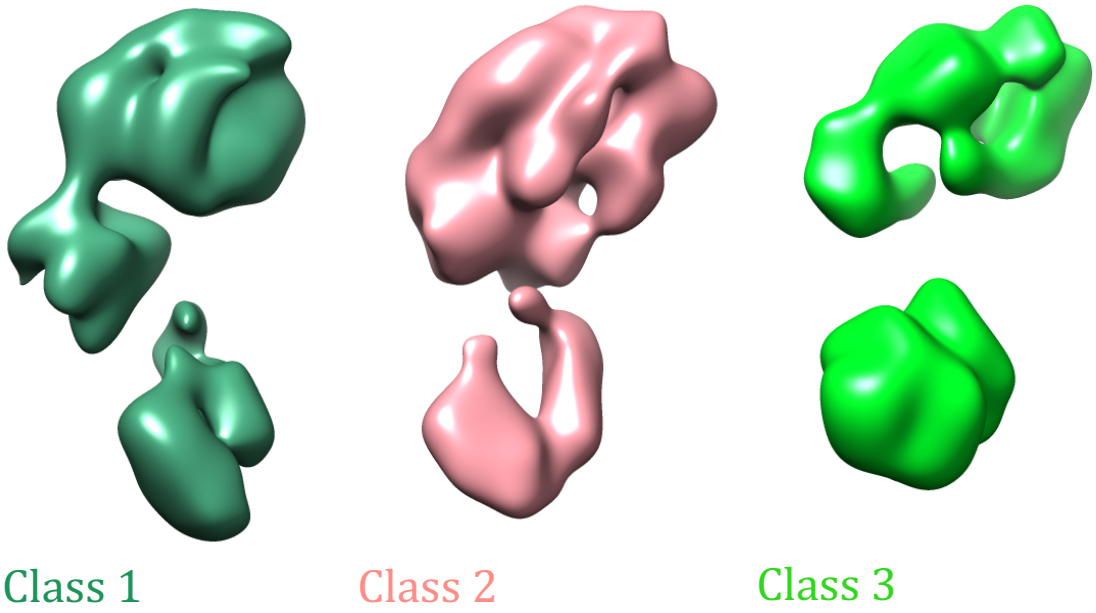


Fig. 4.7 **3D-classification of particles picked without templates, using the CBF3-CEN3 structure as an initial model.** Three 3D-volumes produced by the 3D-classification of the particles picked without templates, using a low pass filtered CBF3-CEN3 atomic model as a reference map.

Class	Class distribution %
1	30.2
2	28.4
3	41.4

Table 4.2 **Class distribution from 3D classification with CBF3-CEN3 model on particles picked without templates.** The table lists percentage of the total particle count occupying each class.

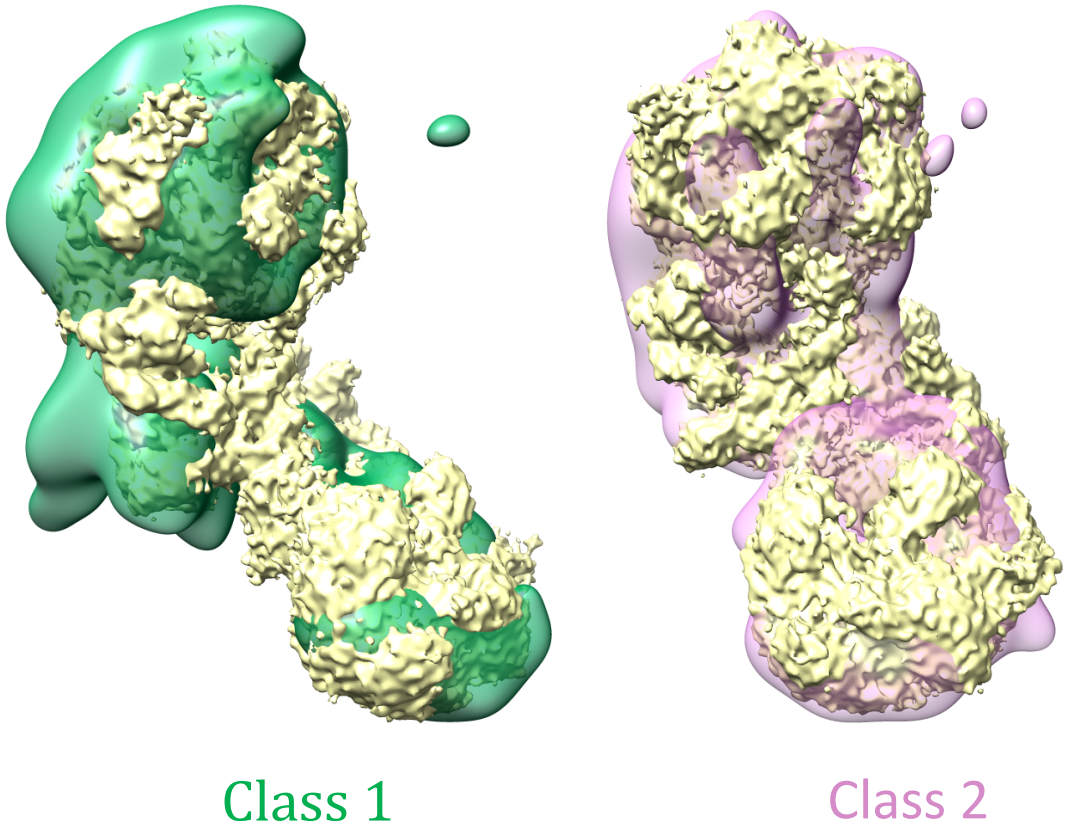


Fig. 4.8 **Map comparison - Class 1 and 2 with CBF3-CEN3.** Class 1 (green) and 2 (pink) from the 3D-classification of particles picked without templates, using CBF3-CEN3 atomic model, was aligned and fitted into the CBF3-CEN3 cryo-EM map (yellow) using the Chimera software. EMDB:0096.

4.3.2 Template based particle picking and image processing

To obtain more particles, Eman was used to generate 800 2D-projections from the cryo-EM map of the CBF3-CEN3 structure, using a script gifted by Luis Miguel. From the 800 projections, 8 different views were selected as templates for particle picking. The templates were low pass filtered to 60 Å, using the Relion-3.0 image handler scripts. The low pass filtered templates were used by Gautomatch v0.56 and particle picking was performed, with a specified particle diameter of 340 Å. Subsequent processing steps were performed using Relion-3.0. Gautomatch selected 81,217 particles which were extracted from the micrographs using a box size of 400 Å. Two rounds of 2D classification were conducted to remove noise as described in Section 2.10.2. The 2D classification for this experiment was performed template-free and using a mask diameter of 360 Å. Both rounds of 2D classification were performed using 200 classes.

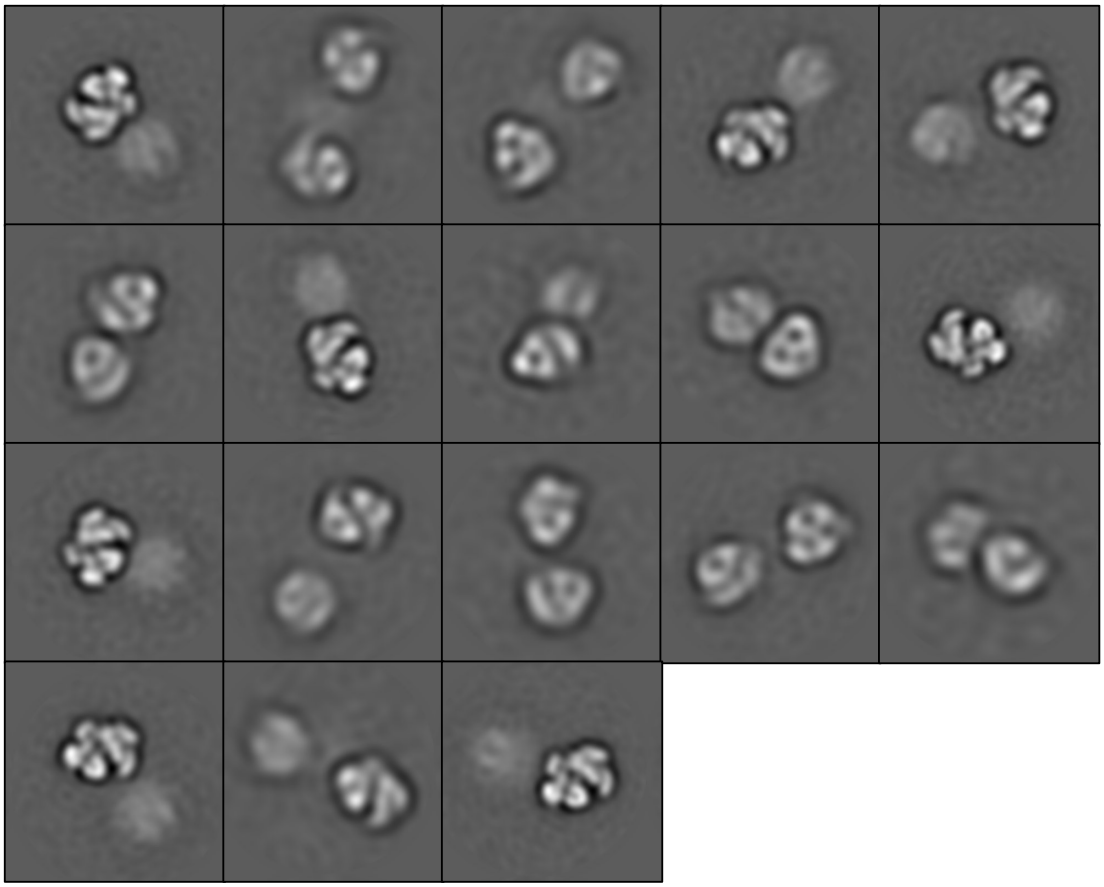


Fig. 4.9 **Selected 2D classes from the template based approach.** The 2D classes of the particles picked with templates,selected for 3D image processing.

The 2D-classes generated (Fig 4.9) were similar in appearance to the 2D-classes of the particles picked without template (Fig 4.3). The majority of the 2D-classes shown contain two particles in close proximity, with one particle more defined than the other. The distance between these particles differs between each classes. Similar to the previous 2D-classification (Fig 4.3), the 2D-classes do not show any connecting density between the particles.

3D classification was performed as described in Section 2.10.2. For comparative purposes with non-template based particle picking strategy (Section 4.3.1), the number of classes was restricted to 3 and the particle diameter was specified as 360 Å. Using the particles selected from 2D classification (Fig 4.9), two Separate 3D classifications were performed, first using the initial model generated with Relion-3.0's SGD algorithm as a reference map and the second using the atomic model of CBF3-CEN3 structure, generated using the Eman pdb2mrc function, which was low pass filtered to 60 Å.

3D classification with SGD initial model generated from the data set

The 3D-classification of particles selected from 2D-classification (Fig 4.9) produced three low resolution 3D-volumes (Fig 4.10), with particle distribution being similar between Class 2 and Class 3 (Table 4.3). The three 3D-maps show the two core CBF3 complexes with density in between representing Ndc10. These classes are similar to Class 1 in Fig 4.4. Of the three 3D-volumes, Class 3 has the most defined connecting density between the two core CBF3 complexes. Class 3 is also the most populated 3D-class, accounting for 31,674 particles (Table 4.3). To confirm the density assignment was correct, Class 3 was manually aligned and fitted into the CBF3-CEN3 cryo-EM map using Chimera (Fig 4.11). The map comparison confirmed that the distance between the core CBF3 complexes was identical between the two maps but the Ndc10 in Class 3 was more flexible and not in the same orientation as the CBF3-CEN3 cryo-EM map (Fig 4.11).

3D-refinement of Class 3

Class 3 (Fig 4.10) had the best defined density for Ndc10 and as a result was selected for 3D-refinement which was performed as described in Section 2.10.2. The particles of Class 3 were selected and a mask diameter of 360 Å was specified. The 3D-map of Class 3 was selected as a reference map for refinement. The 3D-refinement did not improve the quality of the cryo-EM map (Fig 4.12). Both the appearance of the map and the angular distribution show preferential orientation limiting the quality of the

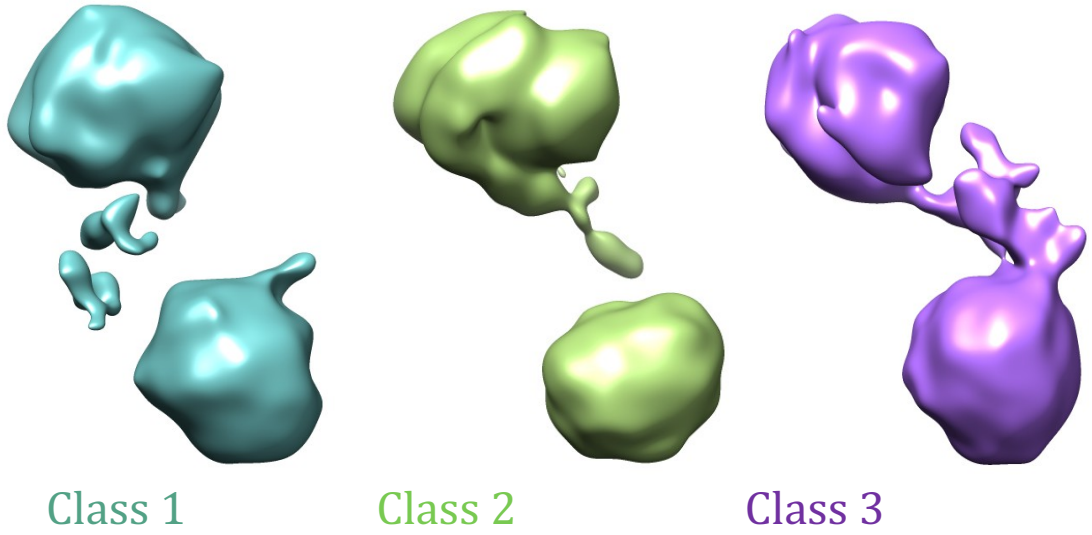


Fig. 4.10 **3D classification with SGD generated initial model.** The three 3D volumes generated after 3D classification with the SGD generated intitial model as a reference map.

Class	Class distribution %
1	25.3
2	35.7
3	39.0

Table 4.3 **Class distribution from the 3D classification with SGD initial model.** The table lists percentage of the total particle count occupying each class.

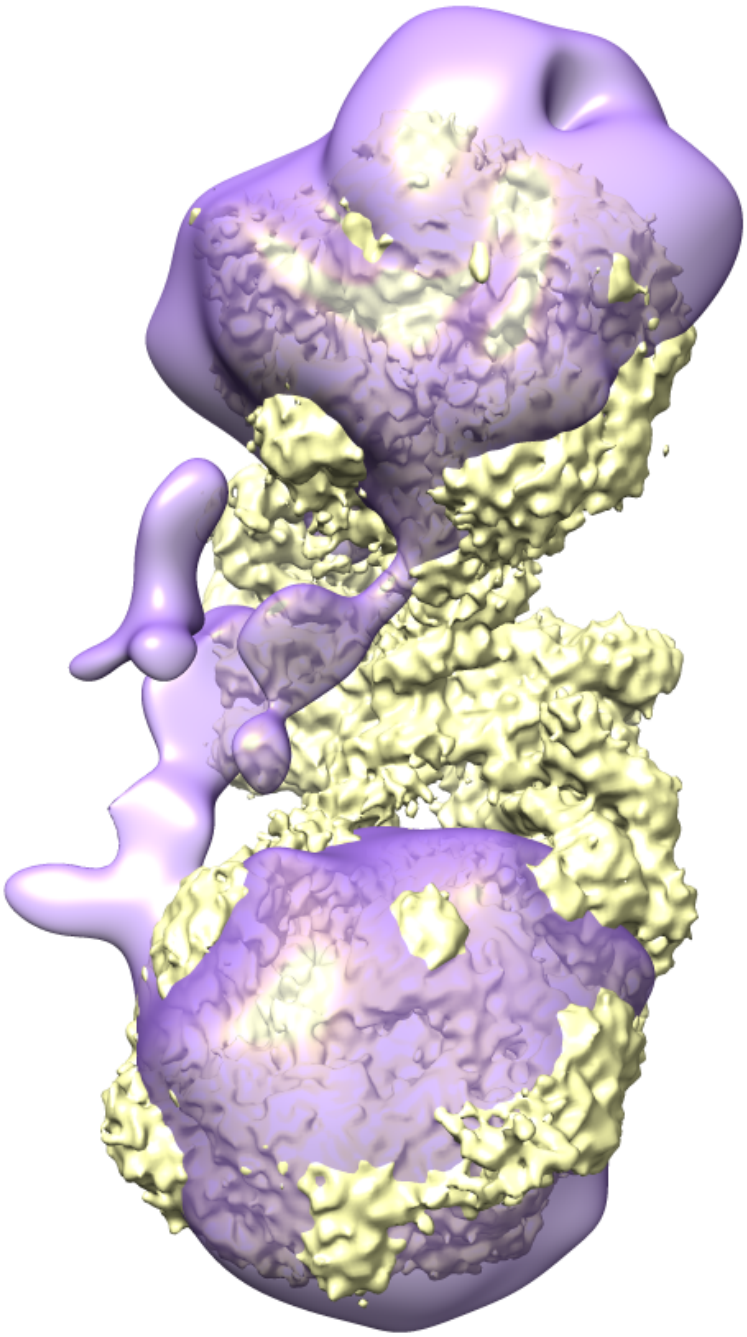


Fig. 4.11 **Map comparison - Class 3 with CBF3-CEN3.** Class 3 (purple) from 3D classification with SGD generated initial model, was fitted into the CBF3-CEN3 cryo-EM map (yellow) using Chimera. EMDB: 0096.

3D-map (Fig 4.12). The densities for the CBF3 core complex and Ndc10 remained ambiguous (Fig 4.12).

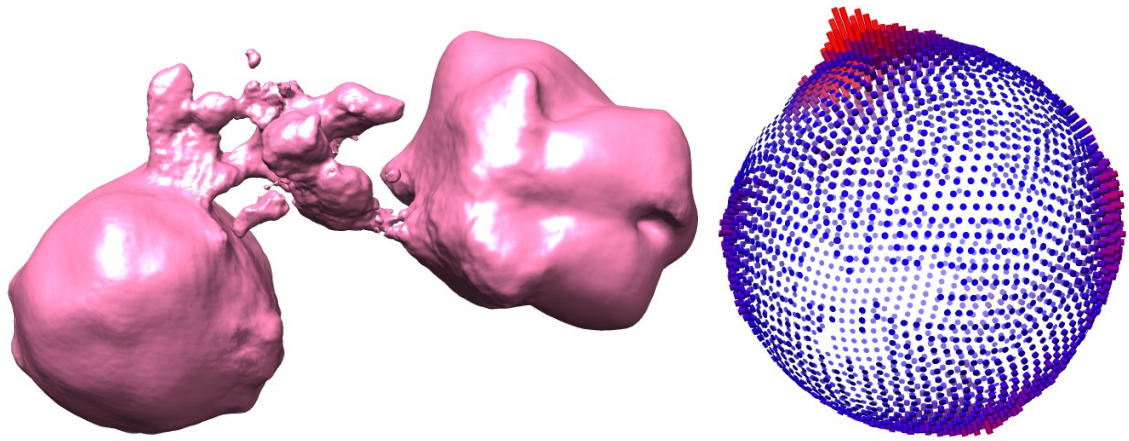


Fig. 4.12 **Refinement of Class 3.** The 3D volumes generated from 3D-refinement using the map of class 3 from 3D-classification as a reference and the angular distribution plot for the refined structure.

3D classification with the CBF3-CEN3 initial model

3D-classification of particles selected from 2D-classification (Fig 4.9) of particles picked using templates, resulted in three low resolution 3D cryo-EM maps (Fig 4.13). Similiar to the 3D-classification conducted with the SGD generated initial model (Fig 4.10), the three classes contain two core CBF3 complexes with a connecting Ndc10. The Ndc10 densities are more defined in the 3D-classes. The class distribution is biased towards Class 3 (Table 4.4), which contains 44,344 particles, which is also the class with the best defined Ndc10. Similiar to previous 3D-classifications, to ensure accurate density assignment, the 3D map of Class 3 was manually aligned and fitted into the 3D cryo-EM map of CBF3-CEN3 using Chimera. Similiar to the map comparisons conducted above, the distance between the CBF3 core complexes were identical between the two maps, with Ndc10 orientation differing (Fig 4.14).

3D-refinement of Class3

Of the three 3D-volumes generated from 3D-classification (Fig 4.13), Class 3 had the highest particle count (Table 4.4) and the best defined Ndc10 and as a result was selected for 3D-refinement. The 3D-refinement was performed as described in Section 2.10.2. The particles from Class 3 were selected, and a mask diameter of 360 Å was

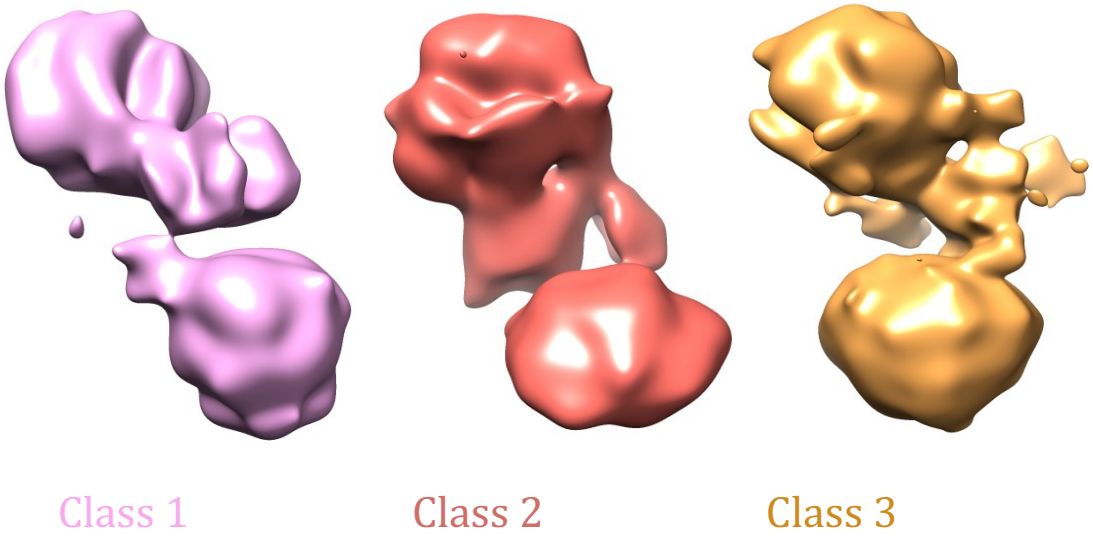


Fig. 4.13 **3D classification with CBF3-CEN3 initial model.** The three 3D volumes generated from 3D classification performed with low pass filtered CBF3-CEN3 structure as a reference map.

Class	Class distribution %
1	24.0
2	21.4
3	54.6

Table 4.4 **Class distribution from the 3D classification with CBF3-CEN3 initial model.** The table lists percentage of the total particle count occupying each class.

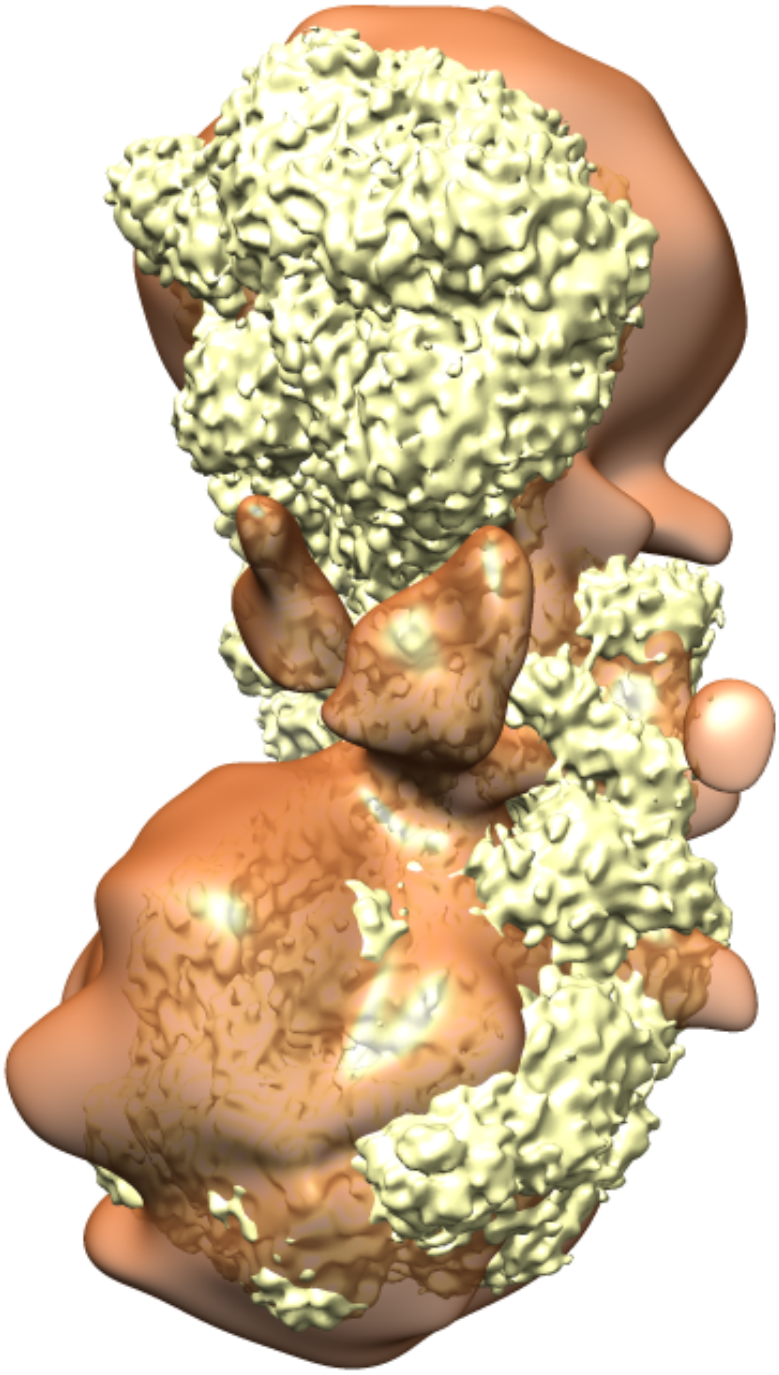


Fig. 4.14 **Map comparison - Class 3 with CBF3-CEN3.** Class 3 (orange) from 3D-classification performed with CBF3-CEN3 atomic model, was fitted into the CBF3-CEN3 cryo-EM map (yellow) using Chimera. EMDB: 0096.

specified for refinement. The 3D-map of Class 3 was selected as the reference map for 3D-refinement. The 3D-refinement showed a good distribution of views present amongst the particles (Fig 4.15). The structural detail of the CBF3 core complexes remained poorly defined, but the Ndc10 density improved. The Ndc10 density suggests a third copy of Ndc10 was present, protruding away from the structure but due to the resolution of the 3D-map this cannot be concluded with confidence.

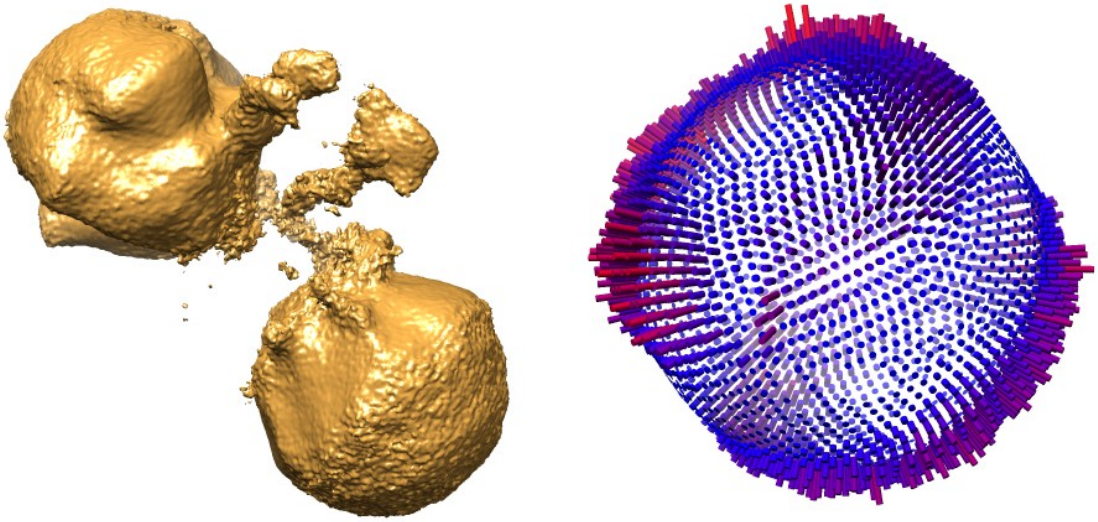


Fig. 4.15 **Refinement of Class 3 from 3D classification performed with CBF3-CEN3 initial model.** Refinement of Class performed on particles, picked using templates, from Class 3 of the 3D-classification performed with the CBF3-CEN3 initial model.

4.4 Monomeric truncated CBF3 complex

The monomeric truncated CBF3 complex is defined as the core CBF3 complex bound to a single copy of Ndc10. Particle picking was performed using Gautomatch v0.56, with no templates and a specified particle diameter of 180 Å. The diameter is based on the size of the truncated CBF3 core complex bound to the DNA binding domain of Ndc10 (Zhang et al., 2018). The coordinates from the particle picking were imported into Relion-3.0. The selected particles were extracted from the micrograph using a 300 Å box, resulting in 224,082 particles. Two rounds of 2D classification were performed

using a mask diameter of 250 Å. This resulted in 10 2D-classes selected, containing 113,209 particles, for 3D-image processing (Fig 4.16).

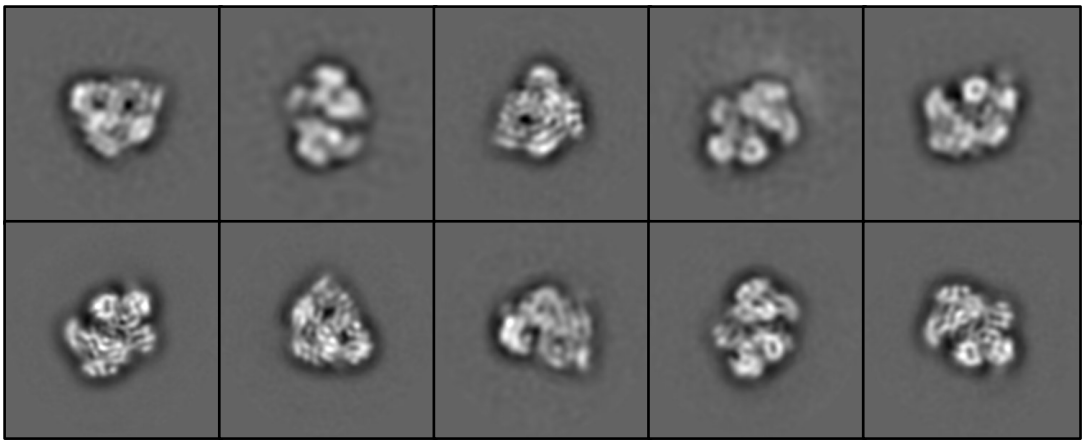


Fig. 4.16 **2D classification of monomeric CBF3.** The 2D-classes of monomeric truncated CBF3 complex,selected for further 3D image processing.

3D-Classification

Initial 3D-classification was performed using the cryo-EM map of the truncated CBF3 complex with the DNA binding domain of Ndc10 (Zhang et al., 2018). The cryo-EM map was low pass fitered to 60 Å, the pixel size was adjusted to 0.85 Å and the box size was adjusted to 300 Å. The particles selected from 2D-classification (Fig 4.16) were classified into 5 classes, with a specified mask diameter of 250 Å. The 3D-maps of the core CBF3 complex were removed, keeping all 3D-classes of the monomeric CBF3 complex. The selected maps underwent 3D-refinement. The refined 3D-map was used to repeat 3D-classification using the same parameters. This procedure was repeated two more times, where each 3D-refined map used as a reference for 3D-classification. This was conducted to remove particles of the CBF3 core complex, and increase the number of particles of the monomeric CBF3 complex.

The final 3D-classification shows five classes with differing densities (Fig 4.17). Class 2 and Class 3 resemble the CBF3 core complex. Class 1 and 4 are the two least populated classes (Table 4.5) which most resemble the monomeric CBF3 complex. Class 5 was ambiguous, with a tail like density projecting from the top of the complex. The appearance of Class 5 could be a result of the the flexibility of Ndc10, resulting in

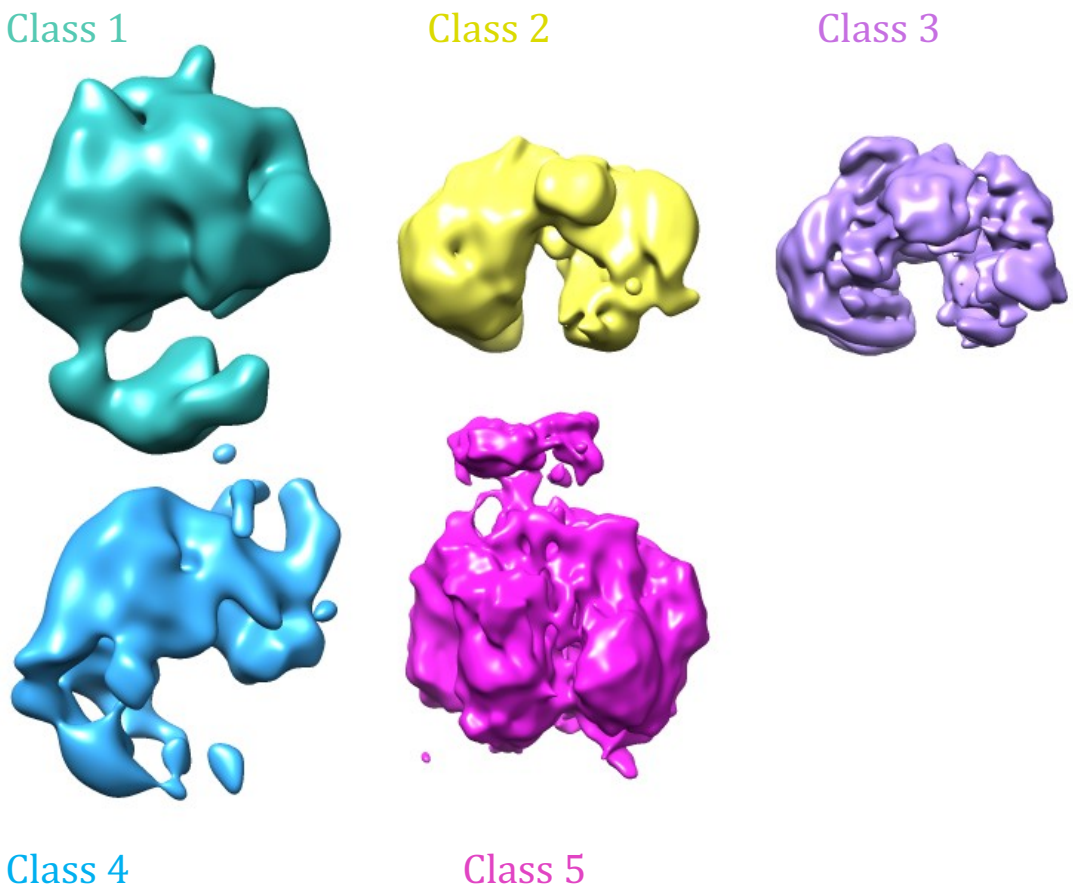


Fig. 4.17 **3D classification of monomeric CBF3.** The five 3D volumes generated from the final 3D classification,after 3 iterations of 3D classification and 3D auto-refinement.

Class	Class distribution %
1	3.8
2	12.0
3	40.0
4	5.4
5	38.8

Table 4.5 **Class distribution from the final 3D classification.** The table lists percentage of the total particle count occupying each class.

the misplacement of the Ndc10 density (Fig 4.17). Class 1, 4 and 5 were selected for 3D-refinement.

3D-Refinement

Particles from Class 1,4 and 5 (Fig 4.17) were selected for 3D-refinement. The 3D-refinement was performed as described in Section 2.10.2, with a mask diameter of 250 Å. The 3D-map after refinement shows the CBF3 core complex, which has the same defined shape as the previously published CBF3 core complex structures (Leber et al., 2018, Zhang et al., 2018). The Ndc10 is well defined with a small interaction site on the CBF3 core complex, resembling the truncated CBF3 complex structure with Ndc10 domains 1-2 (Zhang et al., 2018). The appearance of the map combined with the angular distribution plot (Fig 4.13.B) shows the particles contributing to the map have preferential orientation, with one view dominating the data set. This suggests the orientation of the particles limits the resolution of the 3D-map.

To address the issue of anisotropy, a MATLAB script was gifted by Dr Guilia Zanetti. The script identifies particles of particular views which are over represented in the data set using the particle rotational angle and tilt angles and selects a specified number of particles using the CTF figure of merit. The aim of the script was to reduce the overpresented views, to establish an even angular distribution. This process was performed using Dynamo V1.1.34. The script was used on the particles from the 3D refinement (Fig 4.18), reducing particle count from 54,786 to 45,000. The particle output file was then used to perform 3D-refinement using the new particles. Particles from this refinement underwent additional processing using the script, reducing the particle count from 45,000 to 35,000 and the new particle output file was used to perform 3D refinement again. This process was performed an additional two times, until a final particle count of 22,000 was obtained.

The angular distribution plots after each refinement (Fig 4.19.A) was used to assess the affect of the script on particle distribution. The initial iterations showed little improvement until iteration 4 (Fig 4.19.A). Iteration 4 shows a decrease in the overpopulated angles and more density in the regions which were less represented in iteration 1 - 3 (Fig 4.19.A). Iteration 5 shows further improvement with the red bars, decreasing and the intensity of the blue plots increasing (Fig 4.19.A). The appearance of the map after iteration 5 (Fig 4.19.B) suggests the preferential orientation has less of an effect compared to Fig 4.18. However, the overall map resolution has decreased (Fig 4.19.B).

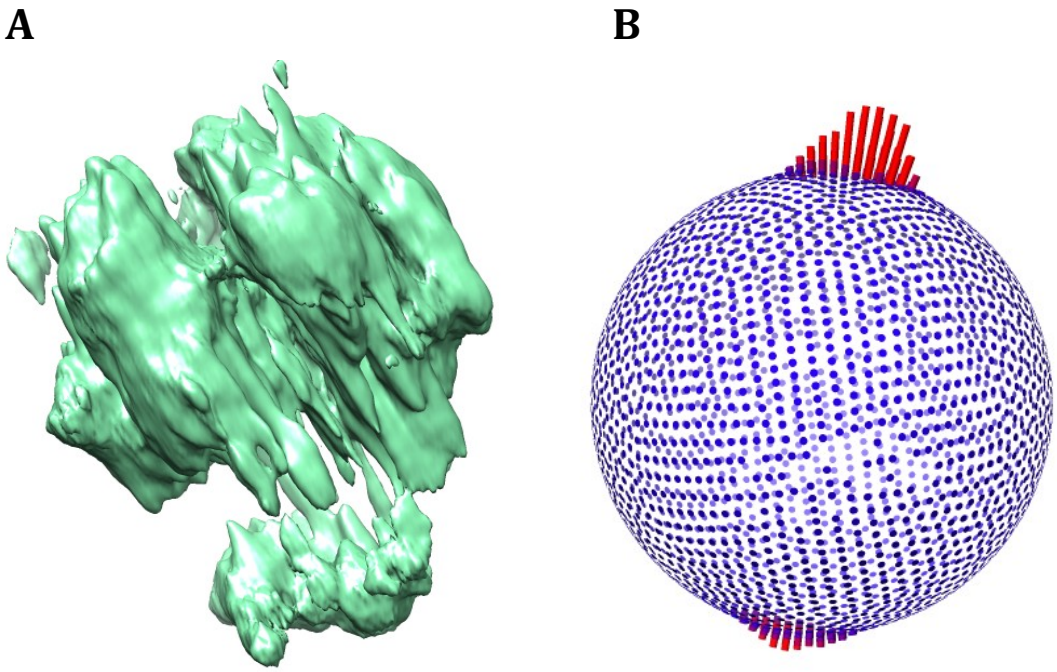


Fig. 4.18 **3D model of monomeric CBF3 complex after refinement** (A) The resulting 3D-refined map after multiple iterations of 3D classification and 3D auto-refinement. (B) angular distribution of the particles contributing to (A).

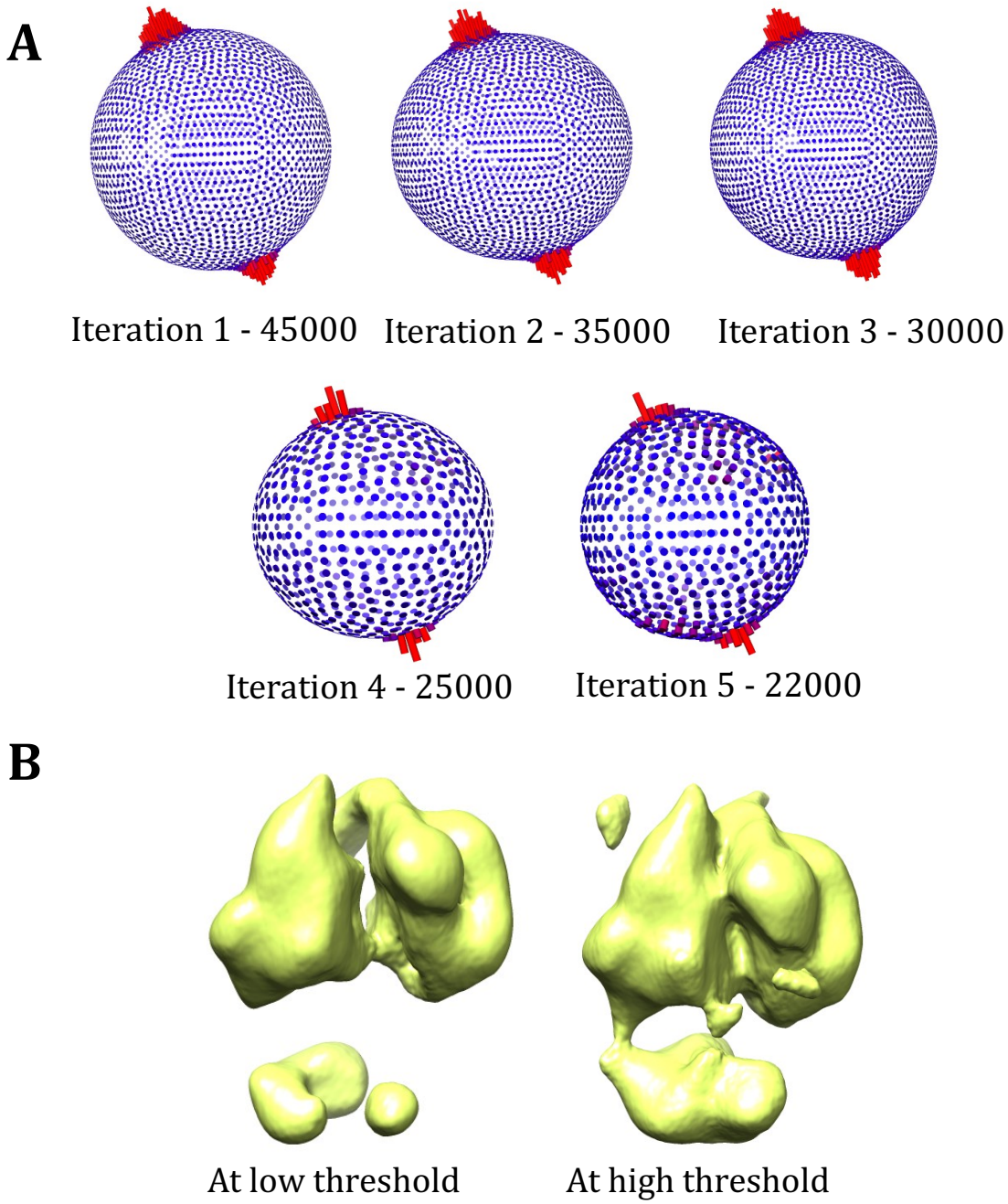


Fig. 4.19 **Fixing anisotropy of the 3D map** (A) The angular distribution after each refinement step. (B) Two 3D-maps displayed with Chimera at low threshold and high threshold.

The decrease in resolution is likely a result of the Ndc10 flexibility and additional views required to obtain a good quality 3D-map is not present in the data set.

4.5 Discussion

Image processing of the CBF3ΔN complex data set has resulted in two 3D-models, a dimeric CBF3 complex and a monomeric CBF3 complex, where CDEIII is not present in either of the resulting models. This was processed from a small population of particles in a heterogenous data set, where the majority of the particles on the micrographs were CBF3ΔN core complex. This was likely a result of the weak interactions between the Ndc10 and the CBF3 core complex which results in the two species readily separating as shown by the chromatogram (Fig 3.10) where the measured elution volume of CBF3ΔN core complex without Ndc10 was significantly higher than the CBF3ΔN complex with Ndc10.

Dimeric CBF3 complex
Using both non-template-based particle picking and template-based particle picking, low resolution 3D-models of the dimeric CBF3 complex were identified. The use of the two approaches to achieve the same model reduced the likelihood of model bias impacting the final result (Henderson, 2013). The obtained 3D-volumes showed two CBF3 core complexes with connecting density between the two core complexes. The obtained 3D-volumes were compared to the published CBF3-CEN3 volumes side by side (Fig 4.20), which showed that all four 3D volumes are of the same size with connecting density between two core complexes and that the separation distance between the CBF3 core complexes were similar (Yan et al., 2018). Thus, these results validate the unexpected bi-lobal CBF3 structural model proposed recently which was not anticipated by other studies.

The obtained dimeric 3D-volume of CBF3 also found that the Ndc10 demonstrated a great degree of flexibility. This led to misalignment, in terms of the position of Ndc10, when the experimental map was superimposed on the published CBF3-CEN3 map. The flexibility also resulted in one of the CBF3 lobes having lower resolution than the other. As the Ndc10 in the map is only composed of Ndc10 domains 1-3, these results suggest that the Ndc10 domains 1-3 provide a flexible link between the two CBF3 core complexes and the comparison between the experimental and published CBF3 maps suggests the centromere DNA restricts the Ndc10 flexibility.

Previous attempts at performing cryo-EM studies of the full CBF3 complex without CEN DNA did not yield a CBF3 complex structure, the CBF3ΔN core complex with the DNA binding domains of Ndc10 did not show a dimeric conformation and the

CBF3-CEN3 study concluded that the dimerisation of the CBF3 complex was induced by the centromeric DNA (Leber et al., 2018, Zhang et al., 2018, Yan et al., 2018). This study obtained a dimeric CBF3 which was not bound to centromeric DNA, however the proportion of dimeric CBF3 was relatively low. Since no population of dimer was observed in the complex with Ndc10 domains 1-2, this study therefore concludes that the dimerisation of the CBF3 complex is initiated by Ndc10 domain 3 (Zhang et al., 2018). The low population of the dimeric complex suggests that the dimerisation initiated by Ndc10 domain 3 is weak and other factors such as binding to centromeric DNA likely enhances the dimerisation.

The CBF3-CEN3 structure and protein-DNA cross-linking experiments have highlighted key interactions sites between Cep3, Ctf13, Ndc10 and the CEN DNA (Espelin et al., 1997, Yan et al., 2018). EMSA studies of Ndc10 have shown only Domains 1-3 play a direct role in DNA interaction (Cho and Harrison, 2012, Perriches and Singleton, 2012). The CBF3-CEN3 structure shows that the Zn_2Cys_6 cluster domains of Cep3 interacted with sites on the CEN DNA but does so in tandem with other interactions between Cep3, Ctf13, Ndc10 and the DNA (Yan et al., 2018). Taken together, this would suggest that the CBF3 complex without the Zn_2Cys_6 cluster domains of Cep3 and domains 4-5 of Ndc10 should be able to bind to CDEIII DNA. Hence, the lack of DNA bound protein complex in this study was surprising. The DNA binding reaction used in this study was similar to the one used in the CBF3-CEN3 study, but centromeric DNA used in the CBF3-CEN3 was greater in length (Yan et al., 2018). The CBF3-CEN3 study also found that the when they studied the CBF3 core complex bound to the DNA binding domain of Ndc10, the resulting structural analysis did not yield a DNA bound protein complex (Yan et al., 2018). Therefore, the lack of DNA bound to the CBF3 in this study could indicate one or a combination of the following: the Zn_2Cys_6 cluster domains of Cep3 are essential for DNA binding, domain 4 and 5 of Ndc10 indirectly contribute to extending the DNA binding surface or the CBF3-DNA binding affinity is affected by the length of the centromeric DNA.

A caveat to comparing the data presented here to the CBF3-CEN3 study is the lack of Zn_2Cys_6 cluster domain of Cep3 in the assembled complex and must be accounted for when drawing comparisons (Yan et al., 2018). It is known when the CCG codon of CDEIII is mutated, the binding site of a Zn_2Cys_6 cluster domain, centromere function is lost, highlighting the fact that the Zn_2Cys_6 cluster domain is essential (Jehn et al., 1991). Future experiments will need to deduce whether the Zn_2Cys_6 cluster domain of Cep3 or Ndc10 domains 4-5 or the combination of both are essential to centromere

binding. To determine whether domain 4 and 5 are essential to DNA binding, the purification and structural analysis needs to be repeated for CBF3ΔN complex with full-length Ndc10 to see whether the complex is able to bind DNA. The role of the Zn_2Cys_6 cluster domain of Cep3 can be probed biochemically and is discussed further in Section 5.7.

Monomeric CBF3 complex

Further analysis of the CBF3ΔN complex data set identified a small sub-population of the total particles represent a monomeric CBF3 complex comprised of a single CBF3ΔN core complex bound to a Ndc10. Initial analysis suggests that the complex closely resembles the CBF3ΔN core complex bound to Ndc10 domains 1-2 structure and the other monomeric core CBF3 complex structures (Zhang et al., 2018, Yan et al., 2018). This indicates that the CBF3 complex could exisit in both monomeric and dimeric form. Additional evaluation of the structure was limited by the low resolution of the 3D-model.

The heterogeneity observed in this data set is consistent with those of other CBF3 structural studies (Leber et al., 2018, Zhang et al., 2018, Yan et al., 2018). These data sets all show that the CBF3 core complex is the most predominant particle, and both this study and the CBF3-CEN3 study have identified a dimeric CBF3 complex, monomeric CBF3 complex and the CBF3 core complex (Leber et al., 2018, Yan et al., 2018). Earlier studies attributed the heterogeneity, or the lack of CBF3 complex, to the Ndc10 forming weak interactions with Ctf13 and Skp1, and that the presence of CEN DNA would strengthen the CBF3 subunit association (Zhang et al., 2018). However, the CBF3-CEN3 study suggests this may not be the case (Yan et al., 2018). The presence of three distinct populations suggests that these CBF3 complexes could represent different stages of the CBF3 assembly pathway and the results presented here suggests that these populations are able to exist without binding to DNA first.

Experimental limitations

The analysis of the obtained 3D-models of the different CBF3 complexes was limited by low resolution and as such future works to improve upon our understanding would need to revolve around improving the resolution. The heterogeneity of the data set was a contributing factor, with only a small population of the particles representing the dimeric and monomeric CBF3ΔN complex. A future consideration could be to use cross-linking to keep the complex intact. This would be a key study, as data without cross linking is already present to ensure the resulting structure is not an

artefact and this would help support the analysis regarding whether the heterogeneity is a result of the distinct states of the CBF3 complex or whether it is a result of the purification and vitrification strategies employed. The resolution of the monomeric CBF3 complex was also limited by preferential orientation. Attempts to reduce the highly populated views did not work, suggesting that the data set lacks an assortment of other views. This issue could be rectified in the future by collecting an additional data set to combine with this data to increase the population of different views to reach a higher resolution. Furthermore, the resolution was also limited by the flexibility of the Ndc10. The flexibility of Ndc10 led to the 2D-class averages of the dimeric CBF3 complexes being poorly centred and the connecting density of the Ndc10 not showing. Subsequent 3D-analysis showed 3D-classification struggled due to misalignment of the Ndc10 with the starting initial models as a result of this flexibility. As such, we can conclude that the dimeric CBF3 complex analysis is limited without the CEN DNA being bound, as the DNA restricts the flexibility of Ndc10. Therefore CBF3 DNA binding limitations must be overcome to proceed the analysis further.

The limitation of CEN DNA not binding to the CBF3ΔN complex has been mentioned a number of times. Whilst this could be due to the reasons discussed above, experimental causes for the lack of DNA binding cannot be dismissed. In the future, the binding reaction needs to be repeated in a different buffer. When Ndc10 domains 1-2 was incubated with both 33 bp CDEIII and 56 bp CDEIII, the resulting gel filtration showed that the protein was unable to bind to either 33 bp or 56 bp CDEIII DNA (See Appendix). This was then rectified by gel filtrating Ndc10 into the 1x EMSA reaction buffer (Section 2.8) and this resulted in the protein binding to DNA (Fig 5.4). Therefore, future experiments need to repeat the binding reaction in the 1x EMSA buffer to see whether this affects DNA binding.

Chapter 5

Results 3: DNA binding activities of the CBF3 core complex and Ndc10

5.1 Introduction

The N-terminal Zn_2Cys_6 cluster domain of Cep3 is the only recognisable DNA binding motif within the CBF3 complex,with strong sequence identity with the Zn_2Cys_6 cluster transcription factors (Schjerling and Holmberg, 1996). The current model for CBF3 function suggests, the Zn_2Cys_6 cluster domain of Cep3 provides DNA sequence specificity, with the DNA binding affinity provided by the rest of Cep3, Ctf13 and Ndc10 (Purvis and Singleton, 2008, Zhang et al., 2018). The N-terminal domain of Cep3 recognises two sequence motifs on CDEIII, the CCG and TGT sites (Purvis and Singleton, 2008). Structural studies of the CBF3 core complex with the N-terminal domain of Cep3, showed that the Zn_2Cys_6 cluster of the Cep3 protomer, situated next to the Ctf13 and Skp1 heterodimer, was ordered and found to interact with an interface formed by the heterodimer (Leber et al., 2018). The N-terminal domain of the other Cep3 protomer was disordered and hence the density was not visible (Leber et al., 2018).

Early work on the CBF3 complex showed the DNA binding activity of the CBF3 complex is dependent on its phosphorylation state (Lechner and Carbon, 1991). Dephos-

phorylated CBF3 complex was unable to bind to the centromere (Lechner and Carbon, 1991). However, recent biochemical analysis of the CBF3 core complex suggests the complex needs to be treated with phosphatase in order to bind to CDEIII DNA (Leber et al., 2018). The site of phosphorylation regulating the DNA binding activity was proposed to be at a N-terminal loop of Skp1 (residues 37 - 64),supported by experiments which show the complex does not require phosphatase treatment to bind DNA when the loop is deleted (Leber et al., 2018). The model proposed suggests the association of Skp1 and Ctf13 with Cep3 locks the core complex in an inactive conformation and dephosphorylation of the Skp1 loop activates the complex for DNA binding (Leber et al., 2018). The experiments localised the effect of Skp1 phosphorylation to the Zn_2Cys_6 cluster domain of Cep3 due to experimental observations (Leber et al., 2018). The first related to the manner in which the zinc cluster domain is docked on to the Ctf13 - Skp1 heterodimer,in the cryo-EM structure, which was not phosphatase treated, suggesting that the Zn_2Cys_6 cluster domain of Cep3 was inactive in the structure (Leber et al., 2018). The second relates to the need for the Zn_2Cys_6 cluster domain of Cep3 for DNA binding (Leber et al., 2018). The EMSA assay conducted showed, with and without phosphatase treatment, DNA binding for the CBF3 core complex without the Zn_2Cys_6 cluster domain of Cep3 (CBF3ΔN core complex), results in no DNA binding (Leber et al., 2018). The latter of the two reasons however was dismissed by a contrasting EMSA assay conducted on the CBF3ΔN core complex, which showed that the complex was capable of binding DNA without the Zn_2Cys_6 cluster domain of Cep3 (Zhang et al., 2018).

The Ndc10 subunit of the CBF3 complex is proposed to provide strong DNA binding affinity to the protein complex (Cho and Harrison, 2012, Perriches and Singleton, 2012). The core DNA binding affinity of Ndc10 comes from domains 1-2, which is referred to as the N-terminal region of Ndc10 (Cho and Harrison, 2012, Perriches and Singleton, 2012, Zhang et al., 2018). Numerous studies have characterised the different properties of the N-terminal region. X-ray crystal structures are available for *S. cerevisiae* Ndc10 domains 1-2 and *K. lactis* Ndc10 domains 1-2 bound to 30 bp CDEIII DNA (Cho and Harrison, 2012, Perriches and Singleton, 2012). The *K. lactis* Ndc10 structures shows Ndc10 associates with the DNA through non-specific interaction with the sugar-phosphate backbone (Cho and Harrison, 2012).

Biochemical studies of *K. lactis* Ndc10 suggests that the dimerisation of Ndc10 is facilitated by Domain 3 of the protein (Cho and Harrison, 2012). Crystal packing analysis of *S. cerevisiae* Ndc10 domains 1-2 shows no biologically relevant dimerisation,

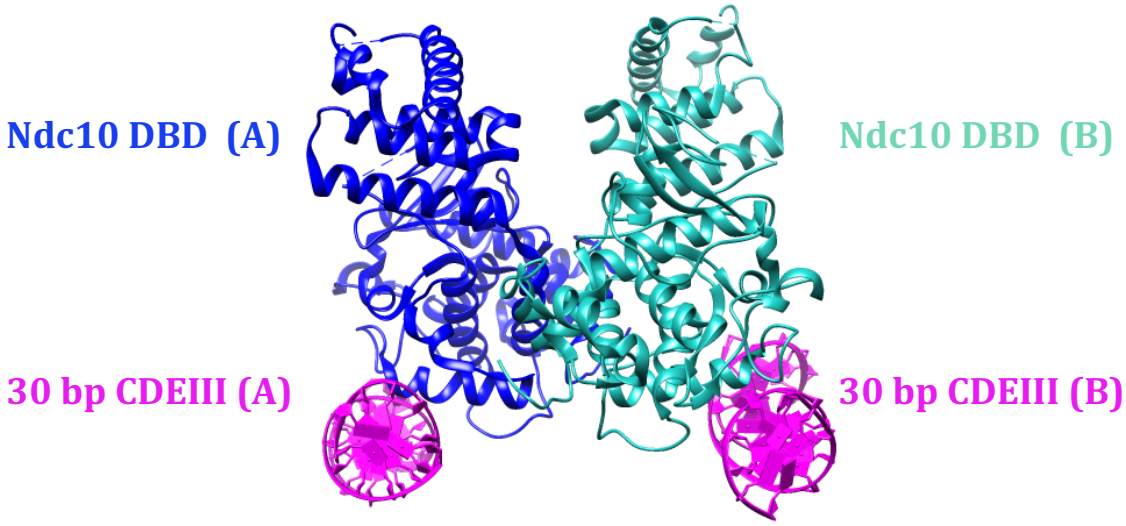


Fig. 5.1 **Dimerisation of Ndc10 DNA Binding Domain.** Crystal structure of *K. lactis* Ndc10 DNA binding domain dimerised binding separate pieces of 30 bp DNA.PDB: 3T79).

which supports biophysical characterisation of these domains which exist as a monomer in solution (Perriches and Singleton, 2012). However, crystal packing analysis of *K. lactis* Ndc10 domains 1-2 reveals an extended interface between two Ndc10 molecules, suggesting that the two neighbouring Ndc10 bound to DNA can dimerise through a small interface of the DNA binding region of Ndc10 (Fig 5.1) (Cho and Harrison, 2012).

This chapter will characterise the CBF3 core complex and its DNA binding function. EMSA experiments were used to determine the effect of (CBF3ΔN core complex) binding to CDEIII at two different lengths: 33 bp and 56 bp. EMSA experiments were also used to determine the contribution of the Zn_2Cys_6 cluster domain of Cep3 to the overall DNA binding affinity of the core complex and how tightly the complex binds to DNA. The role of phosphorylation on CBF3 core complex ability to bind CDEIII was also determined using the EMSA assay. In addition to the CBF3 core complex, this chapter will characterise the Ndc10 DNA binding domains, and whether the N-terminal region of Ndc10 oligomerises in the presence of DNA.

5.2 Sample preparation

CBF3 core complex

In order to prepare the CBF3 core complex to study the role of phosphorylation, the purification of the complex was modified from the purification discussed at the end of Section 3.4. The purification procedure was the same with the exception of two modifications. The 2x Lysis buffer has had the concentration of sodium fluoride adjusted to 100 mM. Both buffers for Nickel affinity column and Heparin chromatography column had been modified to include 50 mM of Sodium fluroide. These modifications ensure phosphatases remain inactive in the lysate and amongst the non-specific proteins bound to each column, which the protein complex could interact with. The gel filtration buffer remained the same.

Ndc10 domains 1-2

The Ndc10 domains 1-2 was prepared as described in Section 3.6. The only modification made was to the gel filtration buffer which was composed of 25 mM Hepes pH 8, 200 mM KCl, 2 mM DTT, 0.02% NP-40 and 10 mM $MgCl_2$. This means that the protein is eluted into the EMSA reaction buffer (Table 2.12) which is required for DNA binding.

5.3 CBF3ΔN core complex binding to CDEIII

Conditions for the EMSA assay were established by Dr Wenjuan Zhang previously (Zhang et al., 2018). The reaction assembly and assay conditions are detailed in Section 2.8. The EMSA assay was performed as described in Section 2.8, using 5' fluorescently labelled 56 bp CDEIII and 33 bp CDEIII. The DNA concentration was kept constant at 3.2 μM. 12.8 μM of protein was used for the protein only control.The protein was titrated in at 3.2 μM, 6.4 μM and 12.8 μM concentrations.

The EMSA assay of CBF3ΔN core complex with 33 bp CDEIII (Fig 5.2.A) is consistent with previous results (Zhang et al., 2018). The EMSA gel (Fig 5.2.A) shows a shift in the position of DNA compared to the DNA only control. The intensity of the DNA band shift increases with increasing protein concentration. When the assay was repeated with 56 bp CDEIII, the results were less clear (Fig 5.2.B). The band shift compared to the DNA only control was less distinct, with strong smearing in each lane. The results suggests that the binding of the CBF3ΔN core complex is diminished with

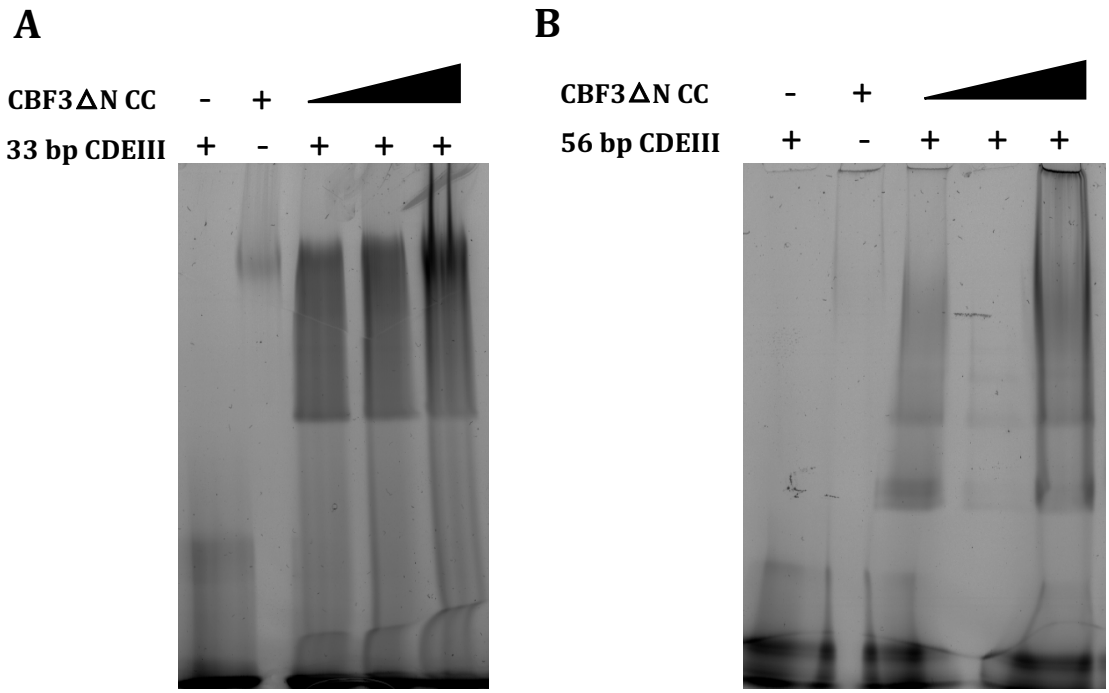


Fig. 5.2 **EMSA assay of the CBF3ΔN core complex with 33 bp and 56 bp CDEIII.** (A) EMSA assay of CBF3ΔN core complex with 56 bp CDEIII. DNA concentration was kept the same at 3.2 μM. The protein complex was titrated in at 3.2 μM, 6.4 μM and 12.8 μM. 12.8 μM of protein complex was used for the protein only control.(B)EMSA assay of the CBF3ΔN core complex with 33 bp CDEIII DNA, using the same DNA and protein concentrations as (A)

increasing length of CDEIII DNA. It is also possible that the 33 bp CDEIII provides a single binding site for the CBF3ΔN core complex but the 56 bp CDEIII can associate with more than 1 core complex but with negative cooperativity. The smearing effect is consistent in both experiments (Fig 5.2) and is indicative of weak affinity and is likely a result of non-specific binding of the CBF3ΔN core complex without the Zn_2Cys_6 cluster domain of Cep3.

5.4 DNA binding activity of the CBF3 core complex

Previously EMSA assays failed to resolve a band shift for the CBF3 core complex when using polyacrylamide gels (Leber et al., 2018). This was also the case for the CBF3ΔN core complex, when using both polyacrylamide based or agarose based gels (Leber et al., 2018). However, our lab successfully used polyacrylamide gels to resolve band shifts for the CBF3ΔN core complex (Zhang et al., 2018). Hence polyacrylamide gels were used to test whether the conditions and parameters established for CBF3ΔN core complex could be used to resolve band shifts for the CBF3 core complex.

The EMSA assay (conditions are detailed in Section 2.8) was conducted using the same reaction parameters and polyacrylamide gels as done for the CBF3ΔN core complex (Section 2.8). The EMSA reaction was done with and without lambda phosphatase treatment, to test whether the phosphorylation determines the CBF3 core complex DNA binding. The protein concentrations used in this assay were lower than the CBF3ΔN core complex because the full-length protein complex does not concentrate as well as the truncated form. The DNA concentration was kept constant at 1.25 μM. The protein was titrated in at 1.25 μM, 2.5 μM and 5 μM. 5 μM of protein complex was used for the protein only controls. 200 units of lambda phosphatase was used for the dephosphorylation reaction.

Titration of the CBF3 core complex at increasing concentrations with 33 bp or 56 bp CDEIII did not result in a strong band shift (Fig 5.3). Neither the 33 bp or the 56 bp CDEIII DNA showed an obvious band shift. The EMSA assay for both 33 bp and 56 bp CDEIII show smearing, which is more intense with 33 bp CDEIII than 56 bp CDEIII. For both 33 bp and 56 bp CDEIII, the CBF3 core complex displays a distinct

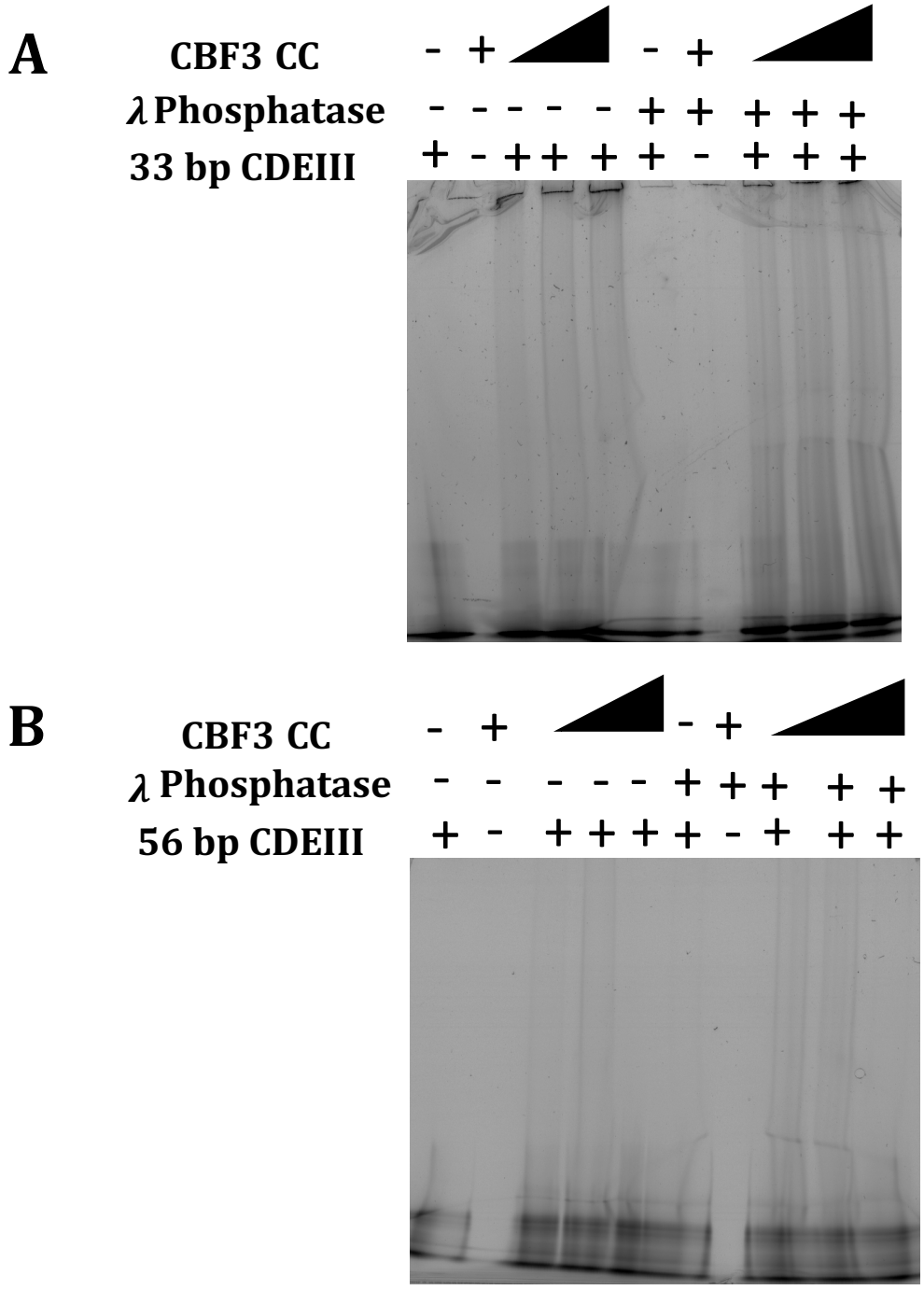


Fig. 5.3 **EMSA assay of the CBF3 core complex with Polyacrylamide gels** (A) EMSA assay of the CBF3 core complex with 33 bp CDEIII. The DNA was at a concentration of 1.25 μM. The protein control was at 5 μM. The protein complex was titrated in at 3.2 μM, 6.4 μM and 12.8 μM. 12.8 μM.The first five lanes are not treated with phosphatase, the last five lanes are treated with phosphatase.(B)EMSA assay of the CBF3ΔN core complex with 56 bp CDEIII DNA, using the same DNA and protein concentrations as (A).

but weak band shift compared to the DNA control, only when the protein complex has been treated with λ phosphatase.

It is unlikely that the weak band shift observed is a result of low protein concentration since this assay has been conducted previously at low protein concentrations but the observation of DNA in the wells (Fig 5.3.A) suggests that a proportion of the CBF3 core complex has not migrated into the gel (Leber et al., 2018). The Zn_2Cys_6 cluster domain of Cep3 may therefore prevent migration into the polyacrylamide gel. This is supported by the fact that the CBF3 core complex without this domain migrates into the polyacrylamide gel, and it is possible to resolve a strong band shift (Fig 5.2). It was concluded that the polyacrylamide performs more poorly at resolving band shifts for the CBF3 core complex than the CBF3ΔN core complex.

In terms of the role of phosphorylation, whilst the weak band shift observed when the CBF3 core complex is treated with λ phosphatase, suggests the complex needs to be dephosphorylated to bind DNA, more work is required to validate the phosphorylation state of the CBF3 core complex before concluding that the band shift is due to dephosphorylation.

5.5 EMSA assay of Ndc10 DNA binding region

In the literature, characterisation of Ndc10 domains 1-2 DNA binding properties have been investigated using EMSA assays conducted with only 30 or 33 bp CDEIII DNA (Cho and Harrison, 2012, Zhang et al., 2018). In order to understand the influence of longer DNA on DNA binding, the first step was to compare the Ndc10 domains 1-2 DNA binding with 33 bp CDEIII to 56 bp CDEIII. The EMSA assay was conducted on Ndc10 domains 1-2 using 5' FAM labelled 33 bp and 56 bp CDEIII. The assay was performed as detailed in Section 2.8. The DNA concentration was kept constant at 3.2 μM and the protein was titrated in at increasing concentrations of 3.2 μM, 6.4 μM and 12.8 μM. 12.8 μM of protein was used for the protein only control.

The EMSA assay conducted using 5' FAM labelled 33 bp CDEIII, shows a primary band shift (Fig 5.4.A), consistent with previous EMSA experiments (Cho and Harrison, 2012, Zhang et al., 2018). There are additional bands present above the primary shift, which can be seen as a smear, but these bands are ambiguous (Fig 5.4.A). When the assay was repeated with 5' labelled 56 bp CDEIII, multiple band shifts were present

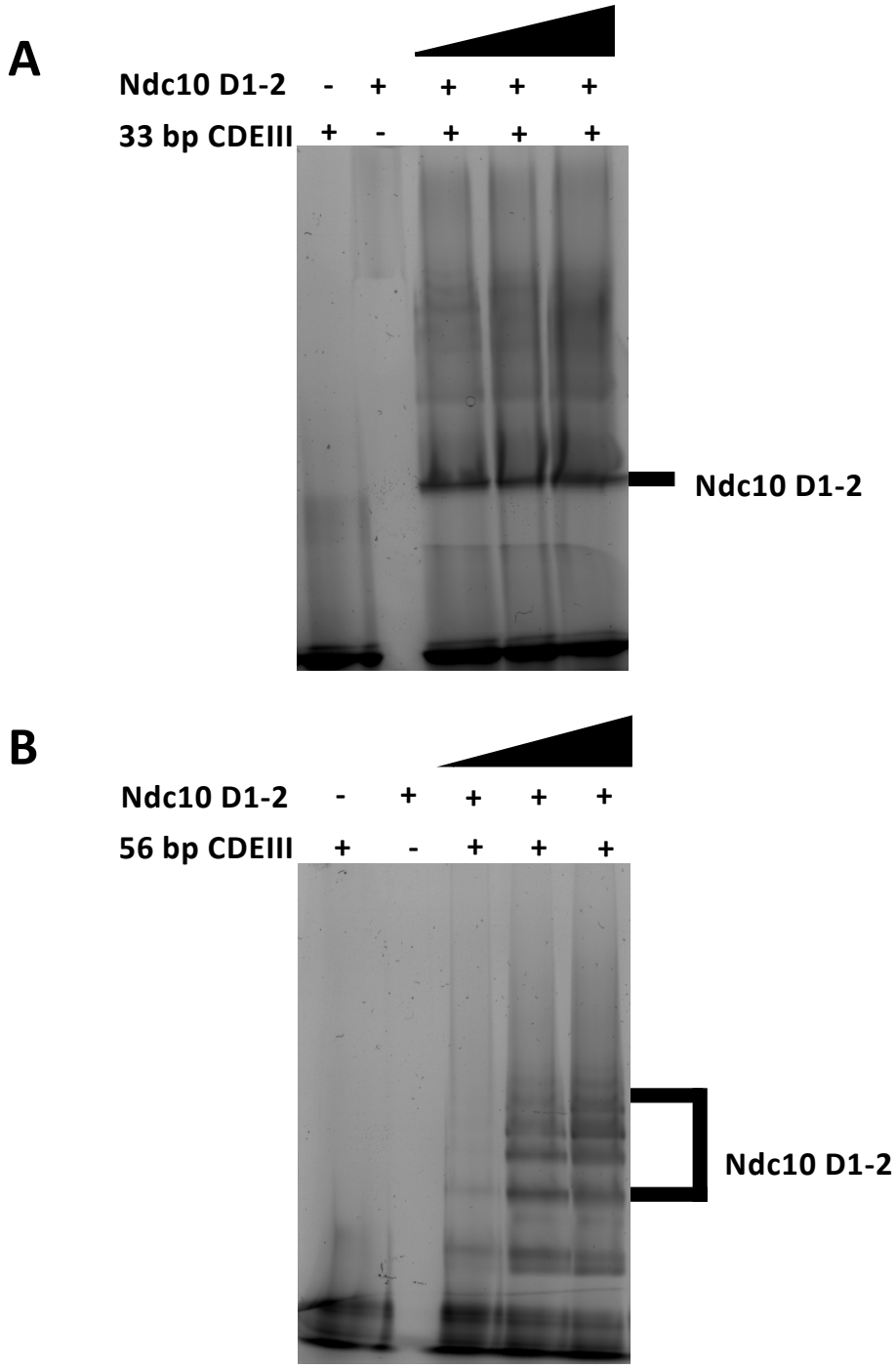


Fig. 5.4 **EMSA assay for Ndc10 DNA Binding Domain.** (A)EMSA assay of Ndc10 domains 1-2 with 33 bp CDEIII. The DNA was at a concentration of 3.2 μM. The protein control was at 12.8 μM. The Ndc10 was titrated at concentrations of 3.2 μM, 6.4 μM and 12.8 μM. (B) The EMSA assay of Ndc10 domains 1-2 with 56 bp CDEIII. The DNA and protein concentrations were the same as (A).

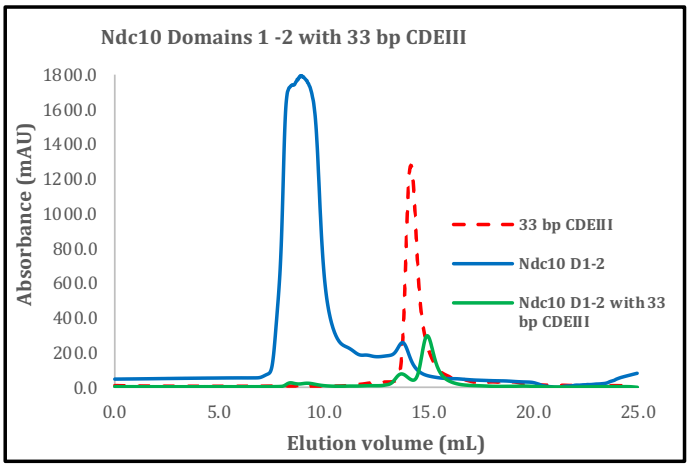
(Fig 5.4.B). Both the number of bands and the intensity of bands increases with increasing amounts of protein. The presence of multiple bands could suggest Ndc10 oligomerisation resulting in differing band shifts and the smearing suggests that either or both the oligomerisation of Ndc10 and the DNA binding by the protein, is weak. However, this cannot conclusively be attributed to either oligomerisation of Ndc10 or DNA binding affinity due to limitations in obtaining the sample. The protein has a tendency to precipitate, highlighted by the large void peak (Fig 3.8). The Ndc10 protein could be aggregating during the experiment resulting in unexpected band shifts.

5.6 DNA induced oligomerisation of Ndc10 domains 1-2

The combination of the crystal packing analysis of *K. lactis* Ndc10 domains 1-2 and the EMSA assay of Ndc10 domains 1-2 with 56 bp CDEIII, generates a new question to address: does Ndc10 domains 1-2 oligomerise in the presence of DNA? To characterise this, a gel filtration experiment was conducted, to compare the elution position of Ndc10 domains 1-2 alone, with 33 bp CDEIII and with 56 bp CDEIII. The rationale behind the experiment was that if the Ndc10 is oligomerising in the presence of DNA, then the protein-DNA complex would elute earlier than the protein alone. The protein was prepared as described in Section 5.2. The Ndc10 domains 1-2 was incubated with a 1.2 molar excess of 33 bp or 56 bp CDEIII at room temperature for 40 minutes and then gel filtrated.

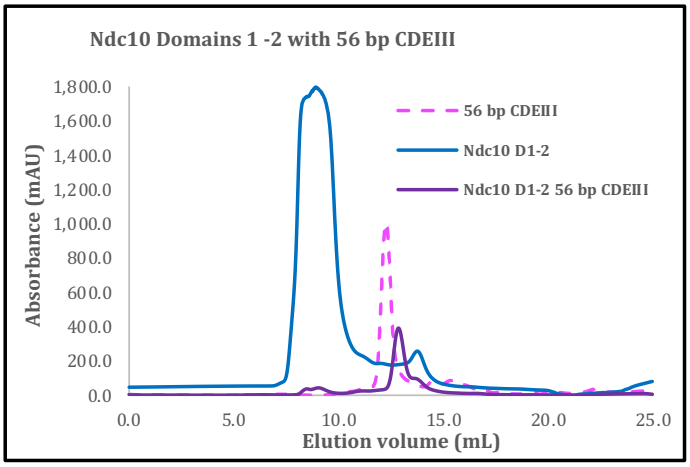
Purification of Ndc10 domains 1-2 alone (Fig 3.8) provided us with data related to the volume at which Ndc10 domains 1-2 elutes from the column, and this acts as a reference for comparison and analysis of Ndc10 domains 1-2 with differing lengths of CDEIII. The purification of Ndc10 domains 1-2 alone also showed a large absorbance measured at the void volume, which was a result of the purification method (Section 3.5). The absorbance peak for Ndc10 domains 1-2 with 33 bp DNA is measured at a higher elution volume than the Ndc10 domains 1-2 alone and the 33 bp DNA (Fig 5.5.A). This suggests that the absorbance is not a result of the DNA eluting alone and SDS-PAGE was used to confirm if the protein was present. A smaller peak is present at the same position as Ndc10 domains 1-2, suggesting a small population of non-bound Ndc10 domains 1-2 was also present (Fig 5.5.A). The elution of the DNA-protein complex occuring later than the Ndc10 alone suggests, the protein becomes more compact with

A



Name	Elution volume Peak (mL)
Ndc10 D1-2	13.75
Ndc10 D1-2 – 33 bp CDEIII	14.93
33 bp CDEIII	14.15

B



Name	Elution volume Peak (mL)
Ndc10 D1-2	13.75
Ndc10 D1-2 – 56 bp CDEIII	12.85
56 bp CDEIII	12.27

Fig. 5.5 **Gel filtration analysis of Ndc10 with CDEIII.** (A)The chromatogram and associated table of the gel filtration analysis for Ndc10 domains 1-2 with 33 bp CDEIII (green). The control peaks of 33 bp CDEIII (dashed red) and Ndc10 domains 1-2 alone (blue) are present for comparison.(B) The chromatogram and the associated table of the gel filtration analysis of Ndc10 domains 1-2 with 56 bp CDEIII.The control peaks of 56 bp CDEIII (dashed pink) and Ndc10 domains 1-2 alone (purple) are present for comparison.

DNA present, as opposed to its extended conformation in the cryo-EM structure when bound to CBF3 (Zhang et al., 2018). The presence of a single peak shift indicates a single population of DNA-protein complex (Fig 5.5.A).

The absorbance peak for Ndc10 domains 1-2 with 56 bp DNA showed the largest measured absorbance occured at an earlier elution volume than the Ndc10 domains 1-2 alone peak (Fig 5.5.B). The measured absorbance was recorded at a higher elution volume than the 56 bp CDEIII DNA alone (Fig 5.5.B). This suggests that the peak is not of only the DNA and SDS-PAGE was used to confirm that protein was present in the recorded peak. There is also a small peak observed at the same position as Ndc10 domain 1 to 2 suggesting a small proportion of Ndc10 has not bound to DNA (Fig 5.5.B). The results from this experiment, suggests the Ndc10 domains 1-2 elutes earlier, an indication of higher molecular weight, but this could also be a result of the fact the protein complex is bound to 56 bp CDEIII which alone elutes earlier than Ndc10 domains 1-2. Therefore it is not possible to conclude if the Ndc10 domains 1-2 is oligomerising. The presence of a single peak shift indicates a single population of DNA-protein complex (Fig 5.5.B).

5.7 Discussion

DNA binding capabilities of the CBF3 Core complex

The EMSA assay of the CBF3ΔN core complex showed that the complex is able to bind to CDEIII DNA in contrast to previous findings (Leber et al., 2018). The CBF3ΔN core complex was able to bind with sufficient affinity to the 33 bp CDEIII DNA resulting in a single band shift, which increases in intensity as the protein concentration increases. However, the binding is less tight with 56 bp CDEIII resulting in a smear on the gel. This band shift was successfully resolved using a polyacrylamide gel.

When the EMSA assay was repeated for the full-length CBF3 core complex, with both 33 bp CDEIII and 56 bp CDEIII DNA, a very weak band shift was observed when the protein complex was treated with λ protein phosphatase (Fig 5.3). Whilst this finding is in agreement to other studies, the limitations of the experiment means no definitive conclusions could be reached (Leber et al., 2018). From this experiment we are unable to tell whether the weak or no band shift is a result of the polyacrylamide gel used in this experiment, an issue observed previously (Leber et al., 2018). Future experiments

could involve repeating this experiment with an agarose based gel to investigate the effect of gel composition on complex electrophoretic mobility. The observation of the weak band shift of the CBF3 core complex cannot be attributed to the activity of the λ protein phosphatase because the phosphorylation state of the complex is not known. Future experiments will require mass spectrometry of the sample before and after the phosphatase treatment in conjunction with the assay to observe if dephosphorylation at specific sites correlates with the band shifts in the assay. Previous studies have identified dephosphorylation at a Skp1 loop (residues 37-64) as being responsible for the CBF3 core complex DNA binding (Leber et al., 2018). This information could be utilised to do additional experiments to help reach a definitive answer. Future experiments could combine phosphomimetic substitution, changing all or individual phosphorylation sites between residues 37-64 on Skp1 to alanine, mimicking the effect of dephosphorylation, and EMSA experiments on the resulting complexes to see whether the CBF3 core complex is able to bind CDEIII DNA. These experiments would enable the field to understand whether the CBF3 complex centromere binding activity is phospho-regulated.

Upon establishing the CBF3 core complex EMSA assay, more can then be done to probe CBF3 – centromere binding mechanism. Future studies will be needed to determine whether the Zn_2Cys_6 cluster domain of Cep3 determines specificity for CDEIII sequence. This can be done be done by conducting the EMSA assay with an unlabelled competitor DNA sequence of no relation to CDEIII and a labelled CDEIII sequence where the label enables the band shift to be observed. If the Zn_2Cys_6 cluster domain of Cep3 determines specificity for CDEIII sequence, the protein complex should not bind to the competitor sequence, and in the presence of both types of DNA there should be no changes to the observed band shift. Changes to the observed band shift would indicate that the protein complex is binding to both pieces of DNA, which would suggest that the DNA binding is non-specific. Control studies in which the same experiments are performed on the CBF3ΔN core complex would provide a reference for the CBF3 core complex data to determine if the Zn_2Cys_6 cluster domain of Cep3 provides specificity for CDEIII sequence.

DNA binding and oligomerisation of Ndc10 domains 1-2

Two approaches were taken to investigate the Ndc10 domains 1-2 DNA binding and oligomerisation properties in the presence of 33 bp and 56 bp CDEIII DNA. The EMSA assay showed, in the presence of 33 bp CDEIII, there was a primary band shift accompanied by a smear whereas with 56 bp CDEIII, there were a number of

distinct bands where their presence was dependent on the concentration of Ndc10 (Fig 5.4.B). One possibility for this outcome is that the Ndc10 oligomerises in the presence of larger CDEIII DNA. To investigate this possibility, a gel filtration assay was conducted. When Ndc10 domains 1-2 was incubated with 33 bp CDEIII there was a notable peak shift which suggests the protein-DNA complex is more compact, but there is no indication that the protein-DNA complex is oligomerising. When the assay was repeated for 56 bp CDEIII there was a single peak shift at an earlier elution volume, but it was not clear whether this was due to oligomerisation. Due to the limitations of the experiments, as a standalone, conclusions regarding Ndc10 domains 1-2 cannot be reached.

Limitations of the EMSA experiment means it cannot be concluded whether the difference in band shifts between 33bp and 56 bp CDEIII experiments is a result of Ndc10 oligomerisation. Both EMSA experiments show strong smearing, which could be a result of sample instability. During the protein purification process, at the gel filtration step, the majority of Ndc10 elutes in the void volume, suggesting that the protein is aggregating. This raises the question regarding the stability of the purified protein, but more specifically, is Ndc10 domains 1-2 aggregating during the EMSA experiment resulting in the observed band shifts? Future works will need to be focused on modifying the purification protocol to enhance protein stability. One way to do this would be to introduce a dialysis step in between ion exchange and gel filtration. At present, the concentration is adjusted from 1M to 200 mM in one step, and such drastic change in salt concentration could promote instability. Dialysis would enable a more gradual decrease to ensure Ndc10 remains soluble. The EMSA experiment can then be repeated with a more stable protein.

There were a number of limitations in the gel filtration experiment. One of the limitations in the experiment was related to protein concentration. In the EMSA assay the protein was in large excess whilst in this experiment, the DNA is in small excess. A subsequent study where the assay is repeated with increasing concentrations of both DNA and protein would be required to potentially gain any insights into Ndc10 oligomerisation. Moreover, a major limitation of this experiment resides with the process of gel filtration. The published structure of CBF3-CEN3, shows within the structure, two Ndc10 DNA binding domains self-associating across a small interface (Fig.1.18) (Yan et al., 2018). Since the protein doesn't associate without DNA, it was concluded that there is low intrinsic affinity for dimerisation (Perriches and Singleton, 2012, Yan et al., 2018). The duration and the resolving nature of gel filtration renders

it unsuitable for examining weak interactions and as a result it may not be the most suitable method for investigating oligomerisation of Ndc10 domains 1-2. A future experiment better suited to addressing this aim could be to use analytical ultracentrifugation, comparing the Ndc10 domains 1-2 sample with 33 bp and 56 bp, to examine a difference in the recorded molecular weight.

Chapter 6

Discussion

The CBF3 complex is an essential component of the budding yeast kinetochore and is a pre-requisite for the recruitment of both other kinetochore components and the specialised Cse4 nucleosome (Sorger et al., 1994, Camahort et al., 2007, Pagliuca et al., 2009, Lang et al., 2018). To date work has been conducted to decipher the functions of the CBF3 complex and to build on this, the focus has now shifted towards building a mechanistic understanding of how the CBF3 complex assembles, and how it carries out its function. This includes understanding how the CBF3 complex recognises and binds to centromeric DNA and how binding to the DNA facilitates the subsequent stages of kinetochore assembly and function. In order to deepen our understanding of CBF3 biology, this thesis established four key aims to address pressing research questions.

To understand where CDEIII binds on the CBF3 core complex and to determine whether the Zn_2Cys_6 cluster domains of Cep3 provide a canonical binding mode, the first aim of this thesis was to purify the CBF3 core complex with bound CDEIII for structural studies, and without CDEIII for EMSA assays, and compare with CBF3ΔN core complex. A purification protocol was successfully devised enabling preliminary EMSA studies. Analysis showed that in the absence of Zn_2Cys_6 cluster domains, CBF3 core complex displayed differing binding affinity, with stronger binding affinity for 33 bp CDEIII than 56 bp CDEIII. The preliminary EMSA analysis on the full-length CBF3 core complex suggests experimental conditions require further optimisation for a more accurate read-out. The combination of the optimisation that was required to derive the final complex with time constraints on the project meant that structural studies of the resulting complex was not possible.

Next, the thesis set out to unravel the potential link between phosphorylation and CBF3-CDEIII DNA binding activity. The aim was to determine the effects of dephosphorylation of the CBF3 complex on CDEIII binding using EMSA assays. Whilst the experiments conducted in this thesis could suggest a link between phosphatase treatment and CBF3 core complex binding to CDEIII DNA, the reliability of the data and subsequent conclusions were limited due to lack of data regarding the phosphorylation state of the complex before and after the treatment. Future experiments to better address this aim are delineated in Section 5.7.

The final set of aims were related to the overall CBF3 architecture and the validation of two existing CBF3 DNA binding models (Zhang et al., 2018, Yan et al., 2018). The thesis aimed to assemble and purify the full CBF3 complex in the presence and absence of CDEIII DNA for structural studies using cryo-EM and perform a comparison study with the recently published CBF3-CEN3 structure in order to validate or dispute those new findings (Yan et al., 2018). A strategy was successfully devised to purify an intact CBF3 complex and although assembling the complex on CDEIII DNA was unsuccessful, the structural analysis identified both a monomeric and dimeric CBF3 complex. The dimeric CBF3 model was consistent with the published CBF3-CEN3 model, indicating a bi-lobal CBF3 dimer (Yan et al., 2018). These data support the novel CBF3 stoichiometry. The structural analysis also shows that the CBF3 complex can dimerise without centromere DNA, with domain 3 of Ndc10 sufficient for this activity, although the resulting dimerisation is weak. The weak dimerisation was reflected in the data set, where only a small population of the total particles were of the dimeric CBF3. The analysis also found that the two CBF3 core complexes are connected by a flexible Ndc10 domains 1-3 linker. This was evident because one of the CBF3 lobes was at a lower resolution than the other, as well as the misalignment of the position of Ndc10 when the experimental map was superimposed on the published CBF3-CEN3 map. Therefore, the structural studies of the CBF3 complex, validated the published CBF3-CEN3 model, showed both monomeric and dimeric CBF3 conformations occur in solution, the dimerisation of the CBF3 complex is initiated by Ndc10 domain 3 and that the two CBF3 core complexes are linked by a flexible Ndc10 domain 3.

By assimilating the data generated from this study with existing CBF3 literature, it is possible to characterise the oligomerisation state of the CBF3 complex and its centromere DNA binding function. The data shows:

(1) The CBF3 complex is dimeric in both the absence and presence of centromere DNA. The cryo-EM analysis showed that without the DNA, a weak dimer forms. The CBF3-CEN3 study showed that the protein complex forms a more robust dimer when bound to centromere DNA (Yan et al., 2018). Previous biochemical analyses have always isolated a CBF3 monomer but did not identify a CBF3 dimer (Lechner and Carbon, 1991, Espelin et al., 1997, Russell et al., 1999). However, the CBF3: DNA cross-linking study did identify an extended CBF3 complex when the length of the centromere DNA was extended (Espelin et al., 1997). This extended CBF3 complex was interpreted to be a monomeric CBF3 bound to a multimer of Ndc10. It is feasible that the data indicated the dimeric CBF3 complex, since the exact composition of the extended complex was not investigated beyond Ndc10 mutation studies and measuring of the molecular weight (Espelin et al., 1997).

(2) Emerging evidence indicates that phosphorylation may play a role in the CBF3 complex binding to the centromere (Leber et al., 2018). The weak evidence from the EMSA studies presented here combined with other EMSA studies suggest the protein complex needs to be dephosphorylated in order to bind to CDEIII DNA (Leber et al., 2018). In terms of whether the phospho-regulation affects the dimerisation of the CBF3 complex prior to DNA binding or whether it directly affects CBF3 complex binding to centromere DNA needs to be investigated in the future.

(3) The Zn_2Cys_6 cluster domains of Cep3 is the determinant of affinity and specificity for CBF3-CDEIII interaction. The EMSA analysis of the CBF3ΔN core complex showed stronger binding affinity for 33 bp than the 56 bp CDEIII, with the core complex only showing measurable affinity for 33 bp CDEIII. EMSA analysis of Ndc10 domains 1-2 also showed stronger affinity than the CBF3ΔN core complex. The structural analysis showed that the truncated CBF3 complex was unable to bind to the 56 bp CDEIII. A number of possible explanations were identified: the lack of Zn_2Cys_6 cluster domains, the lack of domain 4 and 5 of Ndc10, the length of centromere DNA not providing binding affinity or a combination of some or all of these factors. Domains 4 and 5 of Ndc10 have been shown to not contribute to the Ndc10 DNA binding affinity through EMSA studies of Ndc10 and the CBF3-CEN3 structure does not show interactions between these domains and the centromere DNA (Cho and Harrison, 2012, Perriches and Singleton, 2012, Yan et al., 2018). The CBF3-CEN3 structure shows that the 56 bp used in this study matched the DNA sequence which the CBF3 bound in their structure (Yan et al., 2018). Considered together, these data suggest that the lack of DNA binding is most likely due to the absence of the Zn_2Cys_6 cluster domains of Cep3.

The key findings from this thesis, as well as other CBF3 studies, have been used to devise a new updated model for CBF3 assembly and function (Fig 6.1). The Sgt1-Hsp90 chaperone system facilitates the association of the Ctf13-Skp1 heterodimer with the Cep3 homodimer to form the CBF3 core complex. This is followed by the association between the core complex and the Ndc10 monomer to form a monomeric CBF3 complex (Bansal et al., 2004, Catlett and Kaplan, 2006). CBF3 complex dimerisation is next initiated by domain 3 of Ndc10. The next step involves the Zn_2Cys_6 cluster of Cep3 directing the CBF3 complex to the CDEIII region of the point centromere and the binding affinity is enhanced through interactions between CDEIII and Cep3, Ctf13 and Ndc10 (Espelin et al., 1997, Zhang et al., 2018, Yan et al., 2018). The CBF3 complex dimerisation is then stabilised further by the formation of a second dimerisation interface between two Ndc10 DNA binding domains upon interaction with the centromere (Cho and Harrison, 2012, Yan et al., 2018). The binding of the CBF3 complex bends the point centromere to prime the DNA for Cse4 nucleosome interaction (Zhang et al., 2018, Yan et al., 2018). Ndc10 domains 4 and 5 use Scm3 to localise and deposit the Cse4 nucleosome to the point centromere (Cho and Harrison, 2012).

The results of this thesis and the resulting model raises a number of questions to address. Further evidence is required to understand the mechanisms which spatially and temporally regulate CBF3 function. The CBF3 complex may undergo dephosphorylation via a current unknown phosphatase prior to dimerisation or post CBF3 dimerisation (Leber et al., 2018). There is evidence to suggest phosphorylation regulates the ability of CBF3 to bind centromeric DNA, but additional experiments (Section 5. 7) are required to determine whether this is the case (Lechner and Carbon, 1991, Kaplan et al., 1997, Stemmann et al., 2002, Leber et al., 2018). Future work must also be conducted to identify the phosphatase which regulates the CBF3 complex and whether the dephosphorylation event occurs prior to dimerisation or prior to DNA binding.

In addition, the data presented in this thesis shows that the equilibrium between the CBF3 core complex and the CBF3 monomeric and dimeric complexes is favoured towards the more stable CBF3 core complex, since the CBF3 core complex is more populous in the different cryo-EM datasets (Zhang et al., 2018, Leber et al., 2018, Yan et al., 2018). There is a possibility that the equilibrium could be shifted towards the monomeric and dimeric CBF3 complexes in the presence of CEN DNA but due to the caveats of the experiments conducted, it was not possible to determine this. Further research is required to decipher whether additional mechanisms exist in the

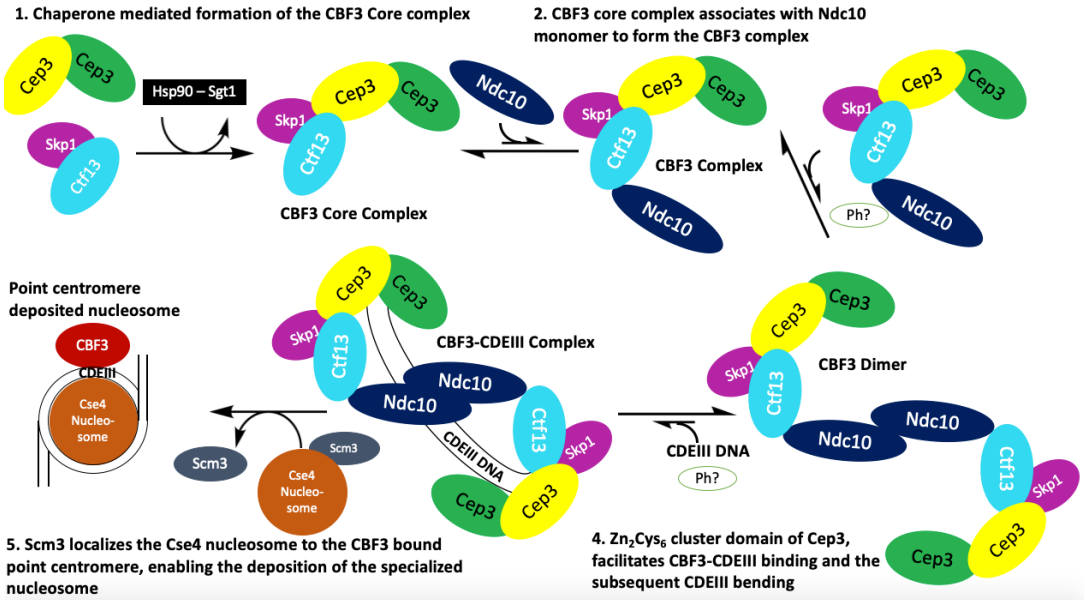


Fig. 6.1 **Revised model of the CBF3 complex formation and function.** The CBF3 core complex formation is facilitated by the Hsp90-Sgt1 chaperone system. The CBF3 core complex binds to Ndc10 to form the CBF3 complex. The equilibrium between the the two complexes are denoted by arrows. The CBF3 dimer formation occurs through CBF3 monomers association via the Ndc10 domain 3. The unknown role of the Phosphatase is denoted in green outlined circle with a question mark. The Zn_2Cys_6 cluster domain of Cep3 directs CDEIII binding and the CBF3 dimer is stabilised by additional Ndc10 N-terminal dimerisation. The Ndc10 domains 4-5 interact with Scm3 to localise and deposit the Cse4 nucleosome to the point centromere.

cell which shifts the equilibrium from the stable core complex to the functional CBF3 monomeric and dimeric complexes, or whether the shift in equilibrium observed in this thesis is a result of the lack of DNA binding. Moreover, little is known as to whether additional proteins regulate CBF3 function post DNA binding. This would require an understanding of the post-translational modification state of the CBF3 complex. Future studies using a combination of mass spectrometry and immunoprecipitation assays could prove to be useful in determining the post-translational modifications found on the CBF3 complex, to identify candidate sites for further investigation.

Moving forward CBF3 studies must focus on building a mechanistic understanding of how CBF3 recruits the Cse4 nucleosome to the point centromere. Since the monomeric CBF3 complex was expected to recruit a single Cse4 nucleosome, the discovery of a dimeric CBF3 complex raises the question: how many Cse4 nucleosomes can the

dimeric CBF3 complex interact with and recruit? Future studies can build on the results in this thesis and other recent studies to express and purify the CBF3 complex and the Cse4 nucleosome to reconstitute a CBF3 complex with centromere DNA wrapped around the nucleosome.

Subsequent research will also need to factor in the CTF19 complex and how CBF3-dependent recruitment of the Cse4 nucleosome and the binding of CTF19 dimer complex affects the CBF3 complex. The unaddressed question relates to how two copies of CTF19, a Cse4 nucleosome and a dimeric CBF3 complex can be accommodated without steric clashes? It has been suggested that CTF19 can disassociate from dimers into monomeric halves (Hinshaw and Harrison, 2019). One possibility is of sequential loading and offloading, where one copy of CTF19 associates with the Cse4 nucleosome, followed by CBF3 disassociating from the centromere, followed by the loading of the second CTF19 in order to accommodate the CTF19 complex. This would explain why the CBF3 does not form a component of the final kinetochore (Akiyoshi et al., 2010, Hinshaw and Harrison, 2019). This is speculative and therefore requires experimental validation. Methods to obtain sufficient quantity of CTF19 complex are published and therefore reconstitution of the CBF3 complex with centromeric DNA and the CTF19 complex is feasible (Hinshaw and Harrison, 2019). This will result in a deeper understanding of how the key components of the inner kinetochore are assembled and regulated and how this influences the assembly and function of the overall kinetochore protein complex.

References

B. Akiyoshi, K. K. Sarangapani, A. F. Powers, C. R. Nelson, S. L. Reichow, H. Arellano-Santoyo, T. Gonen, J. A. Ranish, C. L. Asbury, and S. Biggins. Tension directly stabilizes reconstituted kinetochore-microtubule attachments. *Nature*, 468(7323): 576–579, 2010. ISSN 1476-4687. doi: 10.1038/nature09594.

B. Akiyoshi, C. R. Nelson, and S. Biggins. The aurora B kinase promotes inner and outer kinetochore interactions in budding yeast. *Genetics*, 194(3):785–789, 2013. ISSN 19432631. doi: 10.1534/genetics.113.150839.

D. G. Albertson and J. N. Thomson. The kinetochores of Caenorhabditis elegans. *Chromosoma*, 86(3):409–428, oct 1982. ISSN 1432-0886. doi: 10.1007/BF00292267.

G. M. Alushin, V. H. Ramey, S. Pasqualato, D. A. Ball, N. Grigorieff, A. Musacchio, and E. Nogales. The Ndc80 kinetochore complex forms oligomeric arrays along microtubules. *Nature*, 467(7317):805–10, 2010. ISSN 1476-4687. doi: 10.1038/nature09423.

S. E. Antonarakis, R. Lyle, E. T. Dermitzakis, A. Reymond, and S. Deutsch. Chromosome 21 and down syndrome: from genomics to pathophysiology. *Nature reviews. Genetics*, 5(10):725–38, 2004. ISSN 1471-0056. doi: 10.1038/nrg1448.

P. Aravamudhan, A. A. Goldfarb, and A. P. Joglekar. The kinetochore encodes a mechanical switch to disrupt spindle assembly checkpoint signalling. *Nature Cell Biology*, 17(7):868–879, jun 2015.

C. L. Asbury, D. R. Gestaut, A. F. Powers, A. D. Franck, and T. N. Davis. The Dam1 kinetochore complex harnesses microtubule dynamics to produce force and movement. *Proceedings of the National Academy of Sciences*, 103(26):9873–9878, 2006. ISSN 0027-8424. doi: 10.1073/pnas.0602249103.

C. A. AU - Scarff, M. J. G. AU - Fuller, R. F. AU - Thompson, and M. G. AU - Iadaza. Variations on Negative Stain Electron Microscopy Methods: Tools for Tackling Challenging Systems. *JoVE*, (132):57199, 2018. ISSN 1940-087X. doi: doi:10.3791/57199.

R. E. Baker and K. Rogers. Genetic and genomic analysis of the AT-rich centromere DNA element II of Saccharomyces cerevisiae. *Genetics*, 171(4):1463–1475, 2005. ISSN 00166731. doi: 10.1534/genetics.105.046458.

R. E. Baker, M. Fitzgerald-Hayes, and T. C. O'Brien. Purification of the yeast centromere binding protein CP1 and a mutational analysis of its binding site. *Journal of Biological Chemistry*, 264(18):10843–10850, 1989. ISSN 00219258.

B. E. Bammes, R. H. Rochat, J. Jakana, D.-H. Chen, and W. Chiu. Direct electron detection yields cryo-EM reconstructions at resolutions beyond 3/4 Nyquist frequency. *Journal of Structural Biology*, 177(3):589–601, 2012. ISSN 1047-8477. doi: https://doi.org/10.1016/j.jsb.2012.01.008.

P. K. Bansal, R. Abdulle, and K. Kitagawa. Sgt1 Associates with Hsp90: an Initial Step of Assembly of the Core Kinetochore Complex. *Molecular and Cellular Biology*, 24(18):8069–8079, 2004. ISSN 0270-7306. doi: 10.1128/MCB.24.18.8069-8079.2004.

P. K. Bansal, A. Mishra, A. A. High, R. Abdelle, and K. Kitagawa. Sgt1 dimerization is negatively regulated by protein kinase CK2-mediated phosphorylation at Ser361. *Journal of Biological Chemistry*, 284(28):18692–18698, 2009. ISSN 00219258. doi: 10.1074/jbc.M109.012732.

M. C. Barnhart-Dailey, P. Trivedi, P. T. Stukenberg, and D. R. Foltz. HJURP interaction with the condensin II complex during G1 promotes CENP-A deposition. *Molecular Biology of the Cell*, 28(1):54–64, 2017. ISSN 19394586. doi: 10.1091/mbc.E15-12-0843.

J. J. Bellizzi, P. K. Sorger, and S. C. Harrison. Crystal Structure of the Yeast Inner Kinetochore Subunit Cep3p. *Structure*, 15(11):1422–1430, 2007. ISSN 0969-2126. doi: https://doi.org/10.1016/j.str.2007.09.008.

P. Bernard, J.-F. Maure, J. F. Partridge, S. Genier, J.-P. Javerzat, and R. C. Allshire. Requirement of Heterochromatin for Cohesion at Centromeres. *Science*, 294(5551):2539–2542, 2001. ISSN 0036-8075. doi: 10.1126/science.1064027.

R. Bharadwaj and H. Yu. The spindle checkpoint, aneuploidy, and cancer. *Oncogene*, 23(11):2016–2027, 2004. ISSN 09509232. doi: 10.1038/sj.onc.1207374.

S. Biggins. The composition, functions, and regulation of the budding yeast kinetochore. *Genetics*, 194(4):817–846, 2013. ISSN 00166731. doi: 10.1534/genetics.112.145276.

S. Biggins and A. W. Murray. The budding yeast protein kinase Ipl1/Aurora allows the absence of tension to activate the spindle checkpoint. *Genes and Development*, 15(23):3118–3129, 2001. ISSN 08909369. doi: 10.1101/gad.934801.

B. E. Black, B. E. Black, D. R. Foltz, D. R. Foltz, S. Chakravarthy, S. Chakravarthy, K. Luger, K. Luger, V. L. Woods, V. L. Woods, D. W. Cleveland, and D. W.

Cleveland. Structural determinants for generating centromeric chromatin. *Nature*, 430(6999):578–582, 2004.

B. E. Black, L. E. Jansen, D. R. Foltz, and D. W. Cleveland. Centromere identity, function, and epigenetic propagation across cell divisions. *Cold Spring Harbor Symposia on Quantitative Biology*, 75:403–418, 2010. ISSN 00917451. doi: 10.1101/sqb.2010.75.038.

K. S. Bloom and J. Carbon. Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. *Cell*, 29(2):305–317, jun 1982. ISSN 0092-8674. doi: 10.1016/0092-8674(82)90147-7.

L. J. Bock, C. Pagliuca, N. Kobayashi, R. A. Grove, Y. Oku, K. Shrestha, C. Alfieri, C. Golfieri, A. Oldani, M. Dal Maschio, R. Bermejo, T. R. Hazbun, T. U. Tanaka, and P. De Wulf. Cnn1 inhibits the interactions between the KMN complexes of the yeast kinetochore. *Nature Cell Biology*, 14(6):614–624, 2012. ISSN 14657392. doi: 10.1038/ncb2495.

K. A. Borkovich, F. W. Farrelly, D. B. Finkelstein, J. Taulien, and S. Lindquist. hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures. *Molecular and Cellular Biology*, 9(9):3919–3930, sep 1989. doi: 10.1128/MCB.9.9.3919.

D. Bouck and K. Bloom. The kinetochore protein Ndc10p is required for spindle stability and cytokinesis in yeast. *Proceedings of the National Academy of Sciences*, 102(15):5408–5413, 2005. ISSN 0027-8424. doi: 10.1073/pnas.0405925102.

B.-J. Breitkreutz, C. Stark, T. Reguly, L. Boucher, A. Breitkreutz, M. Livstone, R. Oughtred, D. H. Lackner, J. Bähler, V. Wood, K. Dolinski, and M. Tyers. The BioGRID Interaction Database: 2008 update. *Nucleic Acids Research*, 36(suppl_1): D637–D640, 2007. ISSN 0305-1048. doi: 10.1093/nar/gkm1001.

M. T. Brown. Sequence similarities between the yeast chromosome segregation protein Mif2 and the mammalian centromere protein CENP-C. *Gene*, 160(1):111–116, 1995. ISSN 03781119. doi: 10.1016/0378-1119(95)00163-Z.

M. T. Brown, L. Goetsch, and L. H. Hartwell. MIF2 is required for mitotic spindle integrity during anaphase spindle elongation in Saccharomyces cerevisiae. *Journal of Cell Biology*, 123(2):386–403, 1993. ISSN 00219525.

S. A. Burgess, M. L. Walker, K. Thirumurugan, J. Trinick, and P. J. Knight. Use of negative stain and single-particle image processing to explore dynamic properties of flexible macromolecules. *Journal of Structural Biology*, 147(3):247–258, 2004. ISSN 1047-8477. doi: https://doi.org/10.1016/j.jsb.2004.04.004.

M. Cai and R. W. Davis. Yeast centromere binding protein CBF1, of the helix-loop-helix protein family, is required for chromosome stability and methionine prototrophy. *Cell*, 61(3):437–446, 1990. ISSN 00928674. doi: 10.1016/0092-8674(90)90525-J.

M. J. Cai and R. W. Davis. Purification of a yeast centromere-binding protein that is able to distinguish single base-pair mutations in its recognition site. *Molecular and Cellular biology*, 9(6):2544–50, 1989. ISSN 0270-7306. doi: 10.1128/MCB.9.6.2544. Updated.

R. Camahort, B. Li, L. Florens, S. K. Swanson, M. P. Washburn, and J. L. Gerton. Scm3 Is Essential to Recruit the Histone H3 Variant Cse4 to Centromeres and to Maintain a Functional Kinetochore. *Molecular Cell*, 26(6):853–865, 2007. ISSN 10972765. doi: 10.1016/j.molcel.2007.05.013.

R. Camahort, M. Shivaraju, M. Mattingly, B. Li, S. Nakanishi, D. Zhu, A. Shilatifard, J. L. Workman, and J. L. Gerton. Cse4 Is Part of an Octameric Nucleosome in Budding Yeast. *Molecular Cell*, 35(6):794–805, 2009. ISSN 1097-2765. doi: https://doi.org/10.1016/j.molcel.2009.07.022.

J. Carbon and L. Clarke. Structural and functional analysis of a yeast centromere (CEN3), 1984. ISSN 0269-3518 (Print).

M. G. Catlett and K. B. Kaplan. Sgt1p is a unique co-chaperone that acts as a client adaptor to link Hsp90 to Skp1p. *Journal of Biological Chemistry*, 281(44): 33739–33748, 2006. ISSN 00219258. doi: 10.1074/jbc.M603847200.

I. M. Cheeseman. The Kinetochore. *Cold Spring Harbor Perspectives in Biology*, 6(7): a015826, jul 2014. doi: 10.1101/cshperspect.a015826.

I. M. Cheeseman and A. Desai. Molecular architecture of the kinetochore-microtubule interface. *Nature Reviews Molecular Cell Biology*, 9(1):33–46, 2008. ISSN 1471-0072. doi: 10.1038/nrm2310.

I. M. Cheeseman, S. Anderson, M. Jwa, E. M. Green, J.-S. Kang, J. R. Yates Iii, C. S. M. Chan, D. G. Drubin, and G. Barnes. Phospho-Regulation of Kinetochore-Microtubule Attachments by the Aurora Kinase Ipl1p will require the identification of any remaining kineto- chore proteins. Given the central role that kinetochore-microtubule. *Cell*, 111(2):163–172, 2002.

I. M. Cheeseman, J. S. Chappie, E. M. Wilson-Kubalek, and A. Desai. The Conserved KMN Network Constitutes the Core Microtubule-Binding Site of the Kinetochore. *Cell*, 127(5):983–997, 2006. ISSN 00928674. doi: 10.1016/j.cell.2006.09.039.

Y. Cheng, N. Grigorieff, P. Penczek, and T. Walz. A Primer to Single-Particle Cryo-Electron Microscopy. *Cell*, 161(3):438–449, apr 2015. ISSN 0092-8674. doi: 10.1016/j.cell.2015.03.050.

U.-S. Cho and S. C. Harrison. Recognition of the centromere-specific histone Cse4 by the chaperone Scm3. *Proceedings of the National Academy of Sciences*, 108(23): 9367–9371, 2011. ISSN 0027-8424. doi: 10.1073/pnas.1106389108.

U.-S. Cho and S. C. Harrison. Ndc10 is a platform for inner kinetochore assembly in budding yeast. *Nature structural & molecular biology*, 19(1):48–55, 2012. ISSN 1545-9985. doi: 10.1038/nsmb.2178.

K. Cieśliński and J. Ries. The yeast kinetochore - structural insights from optical microscopy. *Current Opinion in Chemical Biology*, 20(1):1–8, 2014. ISSN 18790402. doi: 10.1016/j.cbpa.2014.03.020.

C. Ciferri, J. De Luca, S. Monzani, K. J. Ferrari, D. Ristic, C. Wyman, H. Stark, J. Kilmartin, E. D. Salmon, and A. Musacchio. Architecture of the Human Ndc80-Hec1 Complex, a Critical Constituent of the Outer Kinetochore. *Journal of Biological Chemistry*, 280(32):29088–29095, aug 2005. doi: 10.1074/jbc.M504070200.

L. Clarke. Centromeres of budding and fission yeasts. *Trends in Genetics*, 6:150–154, jan 1990. ISSN 0168-9525. doi: 10.1016/0168-9525(90)90149-Z.

L. Clarke. Centromeres: proteins, protein complexes, and repeated domains at centromeres of simple eukaryotes. *Current Opinion in Genetics & Development*, 8(2): 212–218, 1998. ISSN 0959437X. doi: 10.1016/S0959-437X(98)80143-3.

L. Clarke and J. Carbon. Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature*, 287(5782):504–509, 1980. ISSN 00280836. doi: 10.1038/287504a0.

H. A. Cole, B. H. Howard, and D. J. Clark. The centromeric nucleosome of budding yeast is perfectly positioned and covers the entire centromere. *Proceedings of the National Academy of Sciences*, 108(31):12687–12692, 2011. ISSN 0027-8424. doi: 10.1073/pnas.1104978108.

C. A. Cooke, D. P. Bazett-Jones, W. C. Earnshaw, and J. B. Rattner. Mapping DNA within the mammalian kinetochore. *Journal of Cell Biology*, 120(5):1083–1091, 1993. ISSN 00219525. doi: 10.1083/jcb.120.5.1083.

K. D. Corbett, C. K. Yip, L. S. Ee, T. Walz, A. Amon, and S. C. Harrison. The monopolin complex crosslinks kinetochore components to regulate chromosome-microtubule attachments. *Cell*, 142(4):556–567, 2010. ISSN 00928674. doi: 10.1016/j.cell.2010.07.017.

J. da Silva Correia, Y. Miranda, N. Leonard, and R. Ulevitch. SGT1 is essential for Nod1 activation. *Proceedings of the National Academy of Sciences*, 104(16): 6764–6769, apr 2007. doi: 10.1073/pnas.0610926104.

A. E. Davies and K. B. Kaplan. Hsp90–Sgt1 and Skp1 target human Mis12 complexes to ensure efficient formation of kinetochore–microtubule binding sites. *The Journal of Cell Biology*, 189(2):261–274, apr 2010. doi: 10.1083/jcb.200910036.

S. De Carlo and H. Stark. Chapter Six - Cryonegative Staining of Macromolecular Assemblies. In G. J. B. T. M. i. E. Jensen, editor, *Cryo-EM Part A Sample Preparation and Data Collection*, volume 481, pages 127–145. Academic Press, 2010. ISBN 0076-6879. doi: https://doi.org/10.1016/S0076-6879(10)81006-7.

P. De Wulf, A. D. McAinsh, and P. K. Sorger. Hierarchical assembly of the budding yeast kinetochore from multiple subcomplexes. *Genes and Development*, 17(23): 2902–2921, 2003. ISSN 08909369. doi: 10.1101/gad.1144403.

O. Díaz-Ingelmo, B. Martínez-García, J. Segura, A. Valdés, and J. Roca. DNA Topology and Global Architecture of Point Centromeres. *Cell Reports*, 13(4):667–677, 2015. ISSN 22111247. doi: 10.1016/j.celrep.2015.09.039.

Y. N. Dimitrova, S. Jenni, R. Valverde, Y. Khin, and S. C. Harrison. Structure of the MIND Complex Defines a Regulatory Focus for Yeast Kinetochore Assembly. *Cell*, 167(4):1014–1027, 2016. ISSN 10974172. doi: 10.1016/j.cell.2016.10.011.

K. F. Doheny, P. K. Sorger, A. A. Hyman, S. Tugendreich, F. Spencer, and P. Hieter. Identification of essential components of the S. cerevisiae kinetochore. *Cell*, 73(4): 761–774, may 1993. ISSN 0092-8674. doi: 10.1016/0092-8674(93)90255-O.

C. Dubacq, R. Guerois, R. Courbeyrette, K. Kitagawa, and C. Mann. Sgt1p Contributes to Cyclic AMP Pathway Activity and Physically Interacts with the Adenylyl Cyclase Cyr1p/Cdc35p in Budding Yeast. *Eukaryotic Cell*, 1(4):568–582, aug 2002. doi: 10.1128/EC.1.4.568-582.2002.

J. Dubochet, M. Adrian, J.-J. Chang, J.-C. Homo, J. Lepault, A. W. McDowall, and P. Schultz. Cryo-electron microscopy of vitrified specimens. *Quarterly Reviews of Biophysics*, 21(2):129–228, 1988. ISSN 0033-5835. doi: DOI:10.1017/S0033583500004297.

E. M. Dunleavy, D. Roche, H. Tagami, N. Lacoste, D. Ray-Gallet, Y. Nakamura, Y. Daigo, Y. Nakatani, and G. Almouzni-Pettinotti. HJURP Is a Cell-Cycle-Dependent Maintenance and Deposition Factor of CENP-A at Centromeres. *Cell*, 137(3):485–497, 2009. ISSN 00928674. doi: 10.1016/j.cell.2009.02.040.

E. H. Egelman. The Current Revolution in Cryo-EM. *Biophysical Journal*, 110(5): 1008–1012, mar 2016. ISSN 0006-3495. doi: 10.1016/j.bpj.2016.02.001.

M. J. Emanuele, W. Lan, M. Jwa, S. A. Miller, C. S. Chan, and P. T. Stukenberg. Aurora B kinase and protein phosphatase 1 have opposing roles in modulating kinetochore assembly. *Journal of Cell Biology*, 181(2):241–254, 2008. ISSN 00219525. doi: 10.1083/jcb.200710019.

H. P. Erickson, A. Klug, H. E. Huxley, and A. Klug. Measurement and compensation of defocusing and aberrations by Fourier processing of electron micrographs. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 261 (837):105–118, may 1971. doi: 10.1098/rstb.1971.0040.

C. W. Espelin, K. B. Kaplan, and P. K. Sorger. Probing the Architecture of a Simple Kinetochore Using DNA-Protein Crosslinking. *The Journal of Cell Biology*, 139(6): 1383–1396, 1997. ISSN 0021-9525. doi: 10.1083/jcb.139.6.1383.

C. W. Espelin, K. T. Simons, S. C. Harrison, and P. K. Sorger. Binding of the Essential Saccharomyces cerevisiae Kinetochore Protein Ndc10p to CDEII. *Molecular Biology of the Cell*, 14(11):4557–4568, sep 2003. ISSN 1059-1524. doi: 10.1091/mbc.e02-08-0533.

J. Fischböck-Halwachs, S. Singh, M. Potocnjak, G. Hagemann, V. Solis-Mezarino, S. Woike, M. Ghodgaonkar-Steger, F. Weissmann, L. D. Gallego, J. Rojas, J. Andreani, A. Köhler, and F. Herzog. The COMA complex interacts with Cse4 and positions Sli15/Ipl1 at the budding yeast inner kinetochore. *eLife*, 8:e42879, 2019. ISSN 2050084X. doi: 10.7554/eLife.42879.

B. Fishel, H. Amstutz, M. Baum, J. Carbon, and L. Clarke. Structural organization and functional analysis of centromeric DNA in the fission yeast Schizosaccharomyces pombe. *Molecular and Cellular Biology*, 8(2):754–763, 1988. ISSN 0270-7306. doi: 10.1128/mcb.8.2.754.

M. Fitzgerald-Hayes, L. Clarke, and J. Carbon. Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. *Cell*, 29(1):235–244, 1982. ISSN 0092-8674. doi: 10.1016/0092-8674(82)90108-8.

D. R. Foltz, L. E. T. Jansen, A. O. Bailey, J. R. Yates, E. A. Bassett, S. Wood, B. E. Black, and D. W. Cleveland. Centromere-Specific Assembly of CENP-A Nucleosomes Is Mediated by HJURP. *Cell*, 137(3):472–484, 2009. ISSN 0092-8674. doi: https://doi.org/10.1016/j.cell.2009.02.039.

T. Fukagawa and W. C. Earnshaw. The centromere: Chromatin foundation for the kinetochore machinery. *Developmental Cell*, 30(5):496–508, 2014. ISSN 18781551. doi: 10.1016/j.devcel.2014.08.016.

M. Funk, R. Niedenthal, D. Mumberg, K. Brinkmann, V. Ro¨nicke, and T. Henkel. Vector systems for heterologous expression of proteins in Saccharomyces cerevisiae. In C. Guthrie and G. R. B. T. M. i. E. Fink, editors, *Guide to Yeast Genetics and Molecular and Cell Biology - Part B*, volume 350, pages 248–257. Academic Press, 2002. ISBN 0076-6879. doi: https://doi.org/10.1016/S0076-6879(02)50967-8.

S. Furuyama and S. Biggins. Centromere identity is specified by a single centromeric nucleosome in budding yeast. *Proceedings of the National Academy of Sciences*, 104 (37):14706–14711, 2007. ISSN 0027-8424. doi: 10.1073/pnas.0706985104.

R. D. Gardner, A. Poddar, C. Yellman, P. A. Tavormina, M. C. Monteagudo, and D. J. Burke. The spindle checkpoint of the yeast Saccharomyces cerevisiae requires kinetochore function and maps to the CBF3 domain. *Genetics*, 157(4):1493–1502, 2001. ISSN 00166731.

K. E. Gascoigne and I. M. Cheeseman. CDK-dependent phosphorylation and nuclear exclusion coordinately control kinetochore assembly state. *Journal of Cell Biology*, 201(1):23–32, 2013. ISSN 00219525. doi: 10.1083/jcb.201301006.

A. Gaudet and M. Fitzgerald-Hayes. Alterations in the adenine-plus-thymine-rich region of CEN3 affect centromere function in Saccharomyces cerevisiae. *Molecular and Cellular Biology*, 7(1):68–75, 1987. ISSN 0270-7306. doi: 10.1128/MCB.7.1.68.

D. R. Gestaut, B. Graczyk, J. Cooper, P. O. Widlund, A. Zelter, L. Wordeman, C. L. Asbury, and T. N. Davis. Phosphoregulation and depolymerization-driven movement of the Dam1 complex do not require ring formation. *Nature Cell Biology*, 10(4): 407–414, 2008. ISSN 14657392. doi: 10.1038/ncb1702.

A. N. Gillis, S. Thomas, S. D. Hansen, and K. B. Kaplan. A novel role for the CBF3 kinetochore-scaffold complex in regulating septin dynamics and cytokinesis. *The Journal of Cell Biology*, 171(5):773–784, 2005. ISSN 0021-9525. doi: 10.1083/jcb.200507017.

R. M. Glaeser. Proteins, interfaces, and cryo-EM grids. *Current Opinion in Colloid & Interface Science*, 34:1–8, 2018. ISSN 1359-0294. doi: https://doi.org/10.1016/j.cocis.2017.12.009.

D. J. Gordon, B. Resio, and D. Pellman. Causes and consequences of aneuploidy in cancer. *Nature reviews. Genetics*, 13(3):189–203, 2012. ISSN 1471-0064. doi: 10.1038/nrg3123.

E. L. Grishchuk, A. K. Efremov, V. A. Volkov, l. S. Spiridonov, N. Gudimchuk, S. Westermann, D. Drubin, G. Barnes, J. R. Mcintosh, and F. l. Ataullakhanov. The Dam1 ring binds microtubules strongly enough to be a processive as well as energy-efficient coupler for chromosome motion. *Proceedings of the National Academy of Sciences*, 105(40):15423–15428, 2008.

Q. Guo, Y. Tao, H. Liu, M. Teng, and X. Li. Structural insights into the role of the Chl4-Iml3 complex in kinetochore assembly. *Acta Crystallographica Section D*, 69 (12):2412–2419, dec 2013. doi: 10.1107/S0907444913022397.

A. Hach, T. Hon, and L. Zhang. The coiled coil dimerization element of the yeast transcriptional activator Hap1, a Gal4 family member, is dispensable for DNA binding but differentially affects transcriptional activation. *Journal of Biological Chemistry*, 275(1):248–254, 2000. ISSN 00219258. doi: 10.1074/jbc.275.1.248.

D. Hanahan. Studies on transformation of Escherichia coli with plasmids. *Journal of Molecular Biology*, 166(4):557–580, 1983. ISSN 0022-2836. doi: https://doi.org/10.1016/S0022-2836(83)80284-8.

A. Hanisch, H. H. W. Silljé, and E. a. Nigg. Timely anaphase onset requires a novel spindle and kinetochore complex comprising Ska1 and Ska2. *The EMBO journal*, 25 (23):5504–5515, 2006. ISSN 0261-4189. doi: 10.1038/sj.emboj.7601426.

J. J. Harrington, G. V. Bokkelen, R. W. Mays, K. Gustashaw, and H. F. Willard. Artificial Microchromosomes. *Nature Genetics*, 15(april):345–355, 1997.

X. He, D. R. Rines, C. W. Espelin, and P. K. Sorger. Molecular analysis of kinetochore-microtubule attachment in budding yeast. *Cell*, 106(2):195–206, 2001. ISSN 00928674. doi: 10.1016/S0092-8674(01)00438-X.

K. A. Hecht, A. F. O'Donnell, and J. L. Brodsky. The proteolytic landscape of the yeast vacuole. *Cellular Logistics*, 4(1):e28023, jan 2014. ISSN null. doi: 10.4161/cl.28023.

J. H. Hegemann and U. N. Fleig. The centromere of budding yeast. *BioEssays*, 15(7): 451–460, 1993. ISSN 0007-1250. doi: 10.1192/bjp.111.479.1009-a.

R. Henderson. Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise. *Proceedings of the National Academy of Sciences*, 110(45):18037–18041, nov 2013. doi: 10.1073/pnas.1314449110.

R. Henderson, A. Sali, M. Baker, B. Carragher, B. Devkota, K. Downing, E. Egelman, Z. Feng, J. Frank, N. Grigorieff, W. Jiang, S. Ludtke, O. Medalia, P. Penczek, P. Rosenthal, M. Rossmann, M. Schmid, G. Schröder, A. Steven, D. Stokes, J. Westbrook, W. Wriggers, H. Yang, J. Young, H. Berman, W. Chiu, G. Kleywegt, and C. Lawson. Outcome of the First Electron Microscopy Validation Task Force Meeting. *Structure*, 20(2):205–214, feb 2012. ISSN 0969-2126. doi: 10.1016/j.str.2011.12.014.

S. Henikoff and T. Furuyama. The unconventional structure of centromeric nucleosomes. *Chromosoma*, 121(4):341–352, aug 2012. ISSN 1432-0886. doi: 10.1007/s00412-012-0372-y.

J. S. P. Heslop-Harrison and T. Schwarzacher. Nucleosomes and centromeric DNA packaging. *Proceedings of the National Academy of Sciences*, 110(50):19974–19975, 2013. ISSN 0027-8424. doi: 10.1073/pnas.1319945110.

S. M. Hinshaw and S. C. Harrison. An Iml3-Chl4 Heterodimer Links the Core Centromere to Factors Required for Accurate Chromosome Segregation. *Cell Reports*, 5 (1):29–36, 2013. ISSN 22111247. doi: 10.1016/j.celrep.2013.08.036.

S. M. Hinshaw and S. C. Harrison. The structure of the ctf19c/ccan from budding yeast. *eLife*, 8:e44239, 2019. ISSN 2050084X. doi: 10.7554/eLife.44239.

A. J. Holland and D. W. Cleveland. Chromoanagenesis and cancer: mechanisms and consequences of localized, complex chromosomal rearrangements. *Nature Medicine*, 18(11):1630–1638, nov 2012.

M. Hong, M. X. Fitzgerald, S. Harper, C. Luo, D. W. Speicher, and R. Marmorstein. Structural Basis for Dimerization in DNA Recognition by Gal4. *Structure*, 16(7): 1019–1026, 2008. ISSN 0969-2126. doi: https://doi.org/10.1016/j.str.2008.03.015.

T.-J. Hong and J.-S. Hahn. Application of SGT1-Hsp90 chaperone complex for soluble expression of NOD1 LRR domain in E. coli. *Biochemical and Biophysical Research Communications*, 478(4):1647–1652, 2016. ISSN 0006-291X. doi: https://doi.org/10.1016/j.bbrc.2016.08.174.

P. Hornung, M. Maier, G. M. Alushin, G. C. Lander, E. Nogales, and S. Westermann. Molecular architecture and connectivity of the budding yeast Mtw1 kinetochore complex. *Journal of Molecular Biology*, 405(2):548–559, 2011. ISSN 00222836. doi: 10.1016/j.jmb.2010.11.012.

P. Hornung, P. Troc, F. Malvezzi, M. Maier, Z. Demianova, T. Zimniak, G. Litos, F. Lampert, A. Schleiffer, M. Brunner, K. Mechtler, F. Herzog, T. C. Marlovits, and S. Westermann. A cooperative mechanism drives budding yeast kinetochore assembly downstream of CENP-A. *Journal of Cell Biology*, 206(4):509–524, 2014. ISSN 15408140. doi: 10.1083/jcb.201403081.

R. Hoser, M. Lichocka, M. Zurczak, J. Hennig, and M. Krzymowska. Emerging role of SGT1 as a regulator of NB-LRRreceptor nucleocytoplasmic partitioning. *Plant Signaling and Behavior*, 9:e28724, 2014. ISSN 15592324. doi: 10.4161/psb.28724.

D. Izawa and J. Pines. The mitotic checkpoint complex binds a second CDC20 to inhibit active APC/C. *Nature*, 517(7536):631–634, 2015. ISSN 14764687. doi: 10.1038/nature13911.

G. J. Jensen. Cryo-EM Part A Sample Preparation and Data Collection. In G. J. B. T. M. i. E. Jensen, editor, *Cryo-EM Part A Sample Preparation and Data Collection*, volume 481, pages xv–xvi. Academic Press, 2010. ISBN 0076-6879. doi: https://doi.org/10.1016/S0076-6879(10)81021-3.

C. Janke, J. Ortíz, T. U. Tanaka, J. Lechner, and E. Schiebel. Four new subunits of the Dam1–Duo1 complex reveal novel functions in sister kinetochore biorientation. *The EMBO Journal*, 21(1-2):181–193, jan 2002. ISSN 0261-4189. doi: 10.1093/emboj/21.1.181.

L. E. T. Jansen, B. E. Black, D. R. Foltz, and D. W. Cleveland. Propagation of centromeric chromatin requires exit from mitosis. *Journal of Cell Biology*, 176(6):795–805, 2007. ISSN 00219525. doi: 10.1083/jcb.200701066.

B. Jehn, R. Niedenthal, and J. H. Hegemann. In vivo analysis of the Saccharomyces cerevisiae centromere CDEIII sequence: requirements for mitotic chromosome segregation. *Molecular and Cellular Biology*, 11(10):5212–5221, 1991.

S. Jenni and S. C. Harrison. Structure of the DASH/Dam1 complex shows its role at the yeast kinetochore-microtubule interface. *Science*, 360(6388):552–558, 2018. ISSN 10959203. doi: 10.1126/science.aar6436.

Z. Ji, H. Gao, and H. Yu. Kinetochore attachment sensed by competitive Mps1 and microtubule binding to Ndc80C. *Science*, 348(6240):1260–1264, 2015. ISSN 0036-8075. doi: 10.1126/science.aaa4029.

A. P. Joglekar. A cell biological perspective on past, present and future investigations of the spindle assembly checkpoint. *Biology*, 5(4):E44, 2016. ISSN 20797737. doi: 10.3390/biology5040044.

A. P. Joglekar, D. C. Bouck, J. N. Molk, K. S. Bloom, and E. D. Salmon. Molecular architecture of a kinetochore-microtubule attachment site. *Nature cell biology*, 8(6): 581–585, 2006. ISSN 1465-7392. doi: 10.1038/ncb1414.

A. P. Joglekar, D. Bouck, K. Finley, X. Liu, Y. Wan, J. Berman, X. He, E. D. Salmon, and K. S. Bloom. Molecular architecture of the kinetochore-microtubule attachment site is conserved between point and regional centromeres. *Journal of Cell Biology*, 181(4):587–594, 2008. ISSN 00219525. doi: 10.1083/jcb.200803027.

A. P. Joglekar, K. Bloom, and E. D. Salmon. In Vivo Protein Architecture of the Eukaryotic Kinetochore with Nanometer Scale Accuracy. *Current Biology*, 19(8): 694–699, 2009. ISSN 09609822. doi: 10.1016/j.cub.2009.02.056.

S. Jones, H. P. Shanahan, H. M. Berman, and J. M. Thornton. Using electrostatic potentials to predict DNA-binding sites on DNA-binding proteins. *Nucleic Acids Research*, 31(24):7189–7198, dec 2003. ISSN 0305-1048. doi: 10.1093/nar/gkg922.

K. B. Kaplan, a. a. Hyman, and P. K. Sorger. Regulating the yeast kinetochore by ubiquitin-dependent degradation and Skp1p-mediated phosphorylation. *Cell*, 91(4): 491–500, 1997. ISSN 0092-8674. doi: S0092-8674(00)80435-3[pii].

F. Katzen. Gateway® recombinational cloning: a biological operating system. *Expert Opinion on Drug Discovery*, 2(4):571–589, 2007. doi: 10.1517/17460441.2.4.571.

S. Kim and H. Yu. Multiple assembly mechanisms anchor the KMN spindle checkpoint platform at human mitotic kinetochores. *Journal of Cell Biology*, 208(2):181–196, 2015. ISSN 15408140. doi: 10.1083/jcb.201407074.

T. Kim, P. Lara-Gonzalez, B. Prevo, F. Meitinger, D. K. Cheerambathur, K. Oegema, and A. Desai. Kinetochores accelerate or delay APC/C activation by directing Cdc20 to opposing fates. *Genes and Development*, 31(11):1089–1094, 2017. ISSN 15495477. doi: 10.1101/gad.302067.117.

K. Kitagawa, D. Skowyra, S. J. Elledge, J. W. Harper, and P. Hieter. SGT1 encodes an essential component of the yeast kinetochore assembly pathway and a novel subunit of the SCF ubiquitin ligase complex. *Molecular Cell*, 4(1):21–33, 1999. ISSN 10972765. doi: 10.1016/S1097-2765(00)80184-7.

K. Kitagawa, R. Abdulle, P. K. Bansal, G. Cagney, S. Fields, and P. Hieter. Requirement of Skp1-Bub1 interaction for kinetochore-mediated activation of the spindle checkpoint. *Molecular Cell*, 11(5):1201–1213, 2003. ISSN 10972765. doi: 10.1016/S1097-2765(03)00145-X.

S. L. Kline, I. M. Cheeseman, T. Hori, T. Fukagawa, and A. Desai. The human Mis12 complex is required for kinetochore assembly and proper chromosome segregation. *Journal of Cell Biology*, 173(1):9–17, 2006. ISSN 00219525. doi: 10.1083/jcb.200509158.

B. Kobe and A. V. Kajava. The leucine-rich repeat as a protein recognition motif. *Current Opinion in Structural Biology*, 11(6):725–732, 2001. ISSN 0959-440X. doi: https://doi.org/10.1016/S0959-440X(01)00266-4.

G. J. Kops, B. A. Weaver, and D. W. Cleveland. On the road to cancer: aneuploidy and the mitotic checkpoint. *Nature Reviews Cancer*, 5(10):773–785, 2005. ISSN 1474-175X. doi: 10.1038/nrc1714.

K. Krassovsky, J. G. Henikoff, and S. Henikoff. Tripartite organization of centromeric chromatin in budding yeast. *Proceedings of the National Academy of Sciences*, 109 (1):243–248, 2012. ISSN 0027-8424. doi: DOI10.1073/pnas.1118898109.

V. Krenn and A. Musacchio. The Aurora B kinase in chromosome bi-orientation and spindle checkpoint signaling. *Frontiers in Oncology*, 5:225, 2015. ISSN 2234943X. doi: 10.3389/fonc.2015.00225.

W. Kühlbrandt. The Resolution Revolution. *Science*, 343(6178):1443–1444, mar 2014.

F. Lampert, P. Hornung, and S. Westermann. The Dam1 complex confers microtubule plus end-tracking activity to the Ndc80 kinetochore complex. *Journal of Cell Biology*, 189(4):641–649, 2010. ISSN 00219525. doi: 10.1083/jcb.200912021.

F. Lampert, C. Mieck, G. M. Alushin, E. Nogales, and S. Westermann. Molecular requirements for the formation of a kinetochore–microtubule interface by Dam1 and Ndc80 complexes. *The Journal of Cell Biology*, 200(1):21–30, jan 2013. doi: 10.1083/jcb.201210091.

J. Lang, A. Barber, and S. Biggins. An assay for de novo kinetochore assembly reveals a key role for the CENP-T pathway in budding yeast. *eLife*, 7:e37819, 2018.

J. Lawrimore, K. S. Bloom, and E. D. Salmon. Point centromeres contain more than a single centromere-specific Cse4 (CENP-A) nucleosome. *Journal of Cell Biology*, 195 (4):573–582, 2011. ISSN 00219525. doi: 10.1083/jcb.201106036.

V. Leber, A. Nans, and M. R. Singleton. Structural basis for assembly of the CBF3 kinetochore complex. *The EMBO Journal*, 37(2):269–281, 2018. ISSN 0261-4189. doi: 10.15252/embj.201798134.

J. Lechner. A zinc finger protein, essential for chromosome segregation, constitutes a putative DNA binding subunit of the Saccharomyces cerevisiae kinetochore complex, Cbf3. *The EMBO Journal*, 13(21):5203–5211, 1994. doi: 10.1002/j.1460-2075.1994.tb06851.x.

J. Lechner and J. Carbon. A 240 kd multisubunit protein complex, CBF3, is a major component of the budding yeast centromere. *Cell*, 64(4):717–725, 1991. ISSN 0092-8674. doi: https://doi.org/10.1016/0092-8674(91)90501-O.

Y. T. Lee, J. Jacob, W. Michowski, M. Nowotny, J. Kuznicki, and W. J. Chazin. Human Sgt1 Binds HSP90 through the CHORD-Sgt1 Domain and Not the Tetratricopeptide Repeat Domain. *Journal of Biological Chemistry*, 279(16):16511–16517, 2004. ISSN 00219258. doi: 10.1074/jbc.M400215200.

X. Li, P. Mooney, S. Zheng, C. R. Booth, M. B. Braunfeld, S. Gubbens, D. A. Agard, and Y. Cheng. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nature Methods*, 10:584, may 2013.

H. Y. Liao and J. Frank. Definition and Estimation of Resolution in Single-Particle Reconstructions. *Structure*, 18(7):768–775, jul 2010. ISSN 0969-2126. doi: 10.1016/j.str.2010.05.008.

S. Lindman, W. F. Xue, O. Szczepankiewicz, M. C. Bauer, H. Nilsson, and S. Linse. Salting the charged surface: pH and salt dependence of protein G B1 stability. *Biophysical Journal*, 90(8):2911–2921, 2006. ISSN 00063495. doi: 10.1529/biophysj.105.071050.

L. B. Lingelbach and K. B. Kaplan. The Interaction between Sgt1p and Skp1p Is Regulated by HSP90 Chaperones and Is Required for Proper CBF3 Assembly. *Molecular and Cellular Biology*, 24(20):8938–8950, 2004. ISSN 0270-7306. doi: 10.1128/mcb.24.20.8938-8950.2004.

D. Liu, M. Vleugel, C. B. Backer, T. Hori, T. Fukagawa, I. M. Cheeseman, and M. A. Lampson. Regulated targeting of protein phosphatase 1 to the outer kinetochore by KNL1 opposes Aurora B kinase. *Journal of Cell Biology*, 188(6):809–820, mar 2010. ISSN 0021-9525. doi: 10.1083/jcb.201001006.

P. S. Maddox, K. Oegema, A. Desai, and I. M. Cheeseman. "Holo"er than thou: Chromosome segregation and kinetochore function in C. elegans. *Chromosome Research*, 12(6):641–653, 2004. ISSN 09673849. doi: 10.1023/B:CHRO.0000036588.42225.2f.

V. Magidson, C. B. O'Connell, J. Lončarek, R. Paul, A. Mogilner, and A. Khodjakov. The spatial arrangement of chromosomes during prometaphase facilitates spindle assembly. *Cell*, 146(4):555–567, 2011. ISSN 00928674. doi: 10.1016/j.cell.2011.07.012.

S. P. Mallick, B. Carragher, C. S. Potter, and D. J. Kriegman. ACE: Automated CTF Estimation. *Ultramicroscopy*, 104(1):8–29, 2005. ISSN 0304-3991. doi: https://doi.org/10.1016/j.ultramic.2005.02.004.

F. Malvezzi, G. Litos, A. Schleiffer, A. Heuck, K. Mechtler, T. Clausen, and S. Westermann. A structural basis for kinetochore recruitment of the Ndc80 complex via two distinct centromere receptors. *The EMBO Journal*, 32(3):409–423, 2013. ISSN 02614189. doi: 10.1038/emboj.2012.356.

D. P. Maskell, X. W. Hu, and M. R. Singleton. Molecular architecture and assembly of the yeast kinetochore MIND complex. *Journal of Cell Biology*, 190(5):823–834, 2010. ISSN 00219525. doi: 10.1083/jcb.201002059.

A. Mayor, F. Martinon, T. De Smedt, V. Pétrilli, and J. Tschopp. A crucial function of SGT1 and HSP90 in inflammasome activity links mammalian and plant innate immune responses. *Nature Immunology*, 8(5):497–503, 2007. ISSN 1529-2916. doi: 10.1038/ni1459.

A. J. McClellan, Y. Xia, A. M. Deutschbauer, R. W. Davis, M. Gerstein, and J. Frydman. Diverse Cellular Functions of the Hsp90 Molecular Chaperone Uncovered Using Systems Approaches. *Cell*, 131(1):121–135, oct 2007. ISSN 0092-8674. doi: 10.1016/j.cell.2007.07.036.

B. F. McEwen, G. K. Chan, B. Zubrowski, M. S. Savoian, M. T. Sauer, and T. J. Yen. CENP-E is essential for reliable bioriented spindle attachment, but chromosome alignment can be achieved via redundant mechanisms in mammalian cells. *Molecular Biology of the Cell*, 12(9):2776–2789, 2001. ISSN 10591524. doi: 10.1091/mbc.12.9.2776.

J. R. McIntosh, E. L. Grishchuk, and R. R. West. Chromosome-Microtubule Interactions During Mitosis. *Annual Review of Cell and Developmental Biology*, 18(1):193–219, 2002. ISSN 1081-0706. doi: 10.1146/annurev.cellbio.18.032002.132412.

K. L. McKinley and I. M. Cheeseman. The molecular basis for centromere identity and function. *Nature Reviews Molecular Cell Biology*, 17:16–29, nov 2016.

G. McMullan, A. R. Faruqi, D. Clare, and R. Henderson. Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy. *Ultramicroscopy*, 147:156–163, 2014. ISSN 0304-3991. doi: https://doi.org/10.1016/j.ultramic.2014.08.002.

G. McMullan, A. R. Faruqi, and R. Henderson. Chapter One - Direct Electron Detectors. In R. A. B. T. M. i. E. Crowther, editor, *The Resolution Revolution: Recent Advances In cryoEM*, volume 579, pages 1–17. Academic Press, 2016. ISBN 0076-6879. doi: https://doi.org/10.1016/bs.mie.2016.05.056.

D. P. Melters, L. V. Paliulis, I. F. Korf, and S. W. L. Chan. Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. *Chromosome Research*, 20(5):579–593, jul 2012. ISSN 1573-6849. doi: 10.1007/s10577-012-9292-1.

P. B. Meluh and D. Koshland. Evidence that the MIF2 gene of Saccharomyces cerevisiae encodes a centromere protein with homology to the mammalian centromere protein CENP-C. *Molecular biology of the cell*, 6(7):793–807, 1995. ISSN 1059-1524. doi: 10.1091/mbc.6.7.793.

P. B. Meluh and D. Koshland. Budding yeast centromere composition and assembly as revealed by in vivo cross-linking. *Genes and Development*, 11(24):3401–3412, 1997. ISSN 08909369. doi: 10.1101/gad.11.24.3401.

C. A. Miller, M. A. Martinat, and L. E. Hyman. Assessment of aryl hydrocarbon receptor complex interactions using pBEVY plasmids: Expression vectors with bi-directional promoters for use in Saccharomyces cerevisiae. *Nucleic Acids Research*, 26(15):3577–3583, 1998. ISSN 03051048. doi: 10.1093/nar/26.15.3577.

J. A. Mindell and N. Grigorieff. Accurate determination of local defocus and specimen tilt in electron microscopy. *Journal of Structural Biology*, 142(3):334–347, 2003. ISSN 1047-8477. doi: https://doi.org/10.1016/S1047-8477(03)00069-8.

J. J. L. Miranda, P. De Wulf, P. K. Sorger, and S. C. Harrison. The yeast DASH complex forms closed rings on microtubules. *Nature structural & molecular biology*, 12(2):138–43, 2005. ISSN 1545-9993. doi: 10.1038/nsmb896.

A. R. Mitchell. The mammalian centromere: Its molecular architecture. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, 372(2):153–162, 1996. ISSN 00275107. doi: 10.1016/S0027-5107(96)00135-2.

T. J. Mitchison and E. D. Salmon. Mitosis: a history of division. *Nature cell biology*, 3 (1):E17–E21, 2001. ISSN 1465-7392. doi: 10.1038/35050656.

G. Mizuguchi, H. Xiao, J. Wisniewski, M. M. Smith, and C. Wu. Nonhistone Scm3 and Histones CenH3-H4 Assemble the Core of Centromere-Specific Nucleosomes. *Cell*, 129(6):1153–1164, 2007. ISSN 00928674. doi: 10.1016/j.cell.2007.04.026.

B. Montpetit, T. R. Hazbun, S. Fields, and P. Hieter. Sumoylation of the budding yeast kinetochore protein Ndc10 is required for Ndc10 spindle localization and regulation of anaphase spindle elongation. *The Journal of Cell Biology*, 174(5):653–663, aug 2006. doi: 10.1083/jcb.200605019.

A. Musacchio and A. Desai. A molecular view of kinetochore assembly and function. *Biology*, 6(1), 2017. ISSN 20797737. doi: 10.3390/biology6010005.

K. Nasmyth. Segregating Sister Genomes: The Molecular Biology of Chromosome Separation. *Science*, 297(5581):559–565, 2002. ISSN 0036-8075. doi: 10.1126/science.1074757.

D. F. Nathan, M. H. Vos, and S. Lindquist. In vivo functions of the Saccharomyces cerevisiae Hsp90 chaperone. *Proceedings of the National Academy of Sciences*, 94 (24):12949–12956, nov 1997. doi: 10.1073/pnas.94.24.12949.

R. Ng and J. Carbon. Mutational and in vitro protein-binding studies on centromere DNA from Saccharomyces cerevisiae. *Molecular and Cellular biology*, 7(12):4522–4534, 1987. ISSN 0270-7306. doi: 10.1128/MCB.7.12.4522.Updated.

V. K. Ngan and L. Clarke. The centromere enhancer mediates centromere activation in Schizosaccharomyces pombe. *Molecular and Cellular Biology*, 17(6):3305–3314, 1997. ISSN 0270-7306. doi: 10.1128/mcb.17.6.3305.

T. Nishino, F. Rago, T. Hori, K. Tomii, I. M. Cheeseman, and T. Fukagawa. CENP-T provides a structural platform for outer kinetochore assembly. *The EMBO Journal*, 32(3):424–436, 2013. ISSN 0261-4189. doi: 10.1038/emboj.2012.348.

J. K. NYBORG and O. B. PEERSEN. That zincing feeling: the effects of EDTA on the behaviour of zinc-binding transcriptional regulators. *Biochemical Journal*, 381 (Pt 3):e3–4, aug 2004. doi: 10.1042/BJ20041096.

M. Ohi, Y. Li, Y. Cheng, and T. Walz. Negative staining and image classification — powerful tools in modern electron microscopy. *Biological Procedures Online*, 6(1): 23–34, 2004. ISSN 1480-9222. doi: 10.1251/bpo70.

E. V. Orlova and H. R. Saibil. Structural Analysis of Macromolecular Assemblies by Electron Microscopy. *Chemical Reviews*, 111(12):7710–7748, dec 2011. ISSN 0009-2665. doi: 10.1021/cr100353t.

J. Ortiz, O. Stemmann, S. Rank, and J. Lechner. A putative protein complex consisting of Ctf19, Mcm21, and Okp1 represents a missing link in the budding yeast kinetochore. *Genes and Development*, 13(9):1140–1155, 1999. ISSN 08909369. doi: 10.1101/gad.13.9.1140.

C. Pagliuca, V. M. Draviam, E. Marco, P. K. Sorger, and P. De Wulf. Roles for the conserved Spc105p/Kre28p complex in kinetochore-microtubule binding and the spindle assembly checkpoint. *PLoS ONE*, 4(10):e7640, 2009. ISSN 19326203. doi: 10.1371/journal.pone.0007640.

B. Panaretou, G. Siligardi, P. Meyer, A. Maloney, J. K. Sullivan, S. Singh, S. H. Millson, P. A. Clarke, S. Naaby-Hansen, R. Stein, R. Cramer, M. Mollapour, P. Workman, P. W. Piper, L. H. Pearl, and C. Prodromou. Activation of the ATPase Activity of Hsp90 by the Stress-Regulated Cochaperone Aha1. *Molecular Cell*, 10(6):1307–1318, dec 2002. ISSN 1097-2765. doi: 10.1016/S1097-2765(02)00785-2.

F. Pangilinan and F. Spencer. Abnormal kinetochore structure activates the spindle assembly checkpoint in budding yeast. *Molecular Biology of the Cell*, 7(8):1195–1208, 1996. ISSN 10591524. doi: 10.1091/mbc.7.8.1195.

G. Pekgöz Altunkaya, F. Malvezzi, Z. Demianova, T. Zimniak, G. Litos, F. Weissmann, K. Mechtler, F. Herzog, and S. Westermann. CCAN Assembly Configures Composite Binding Interfaces to Promote Cross-Linking of Ndc80 Complexes at the Kinetochore. *Current Biology*, 26(17):2370–2378, 2016. ISSN 09609822. doi: 10.1016/j.cub.2016.07.005.

P. A. Penczek. Chapter Two - Image Restoration in Cryo-Electron Microscopy. In G. J. B. T. M. i. E. Jensen, editor, *Cryo-EM, Part B: 3-D Reconstruction*, volume 482, pages 35–72. Academic Press, 2010. ISBN 0076-6879. doi: https://doi.org/10.1016/S0076-6879(10)82002-6.

Y. Peng, C. C. L. Wong, Y. Nakajima, R. G. Tyers, A. S. Sarkeshik, J. Yates, D. G. Drubin, and G. Barnes. Overlapping kinetochore targets of CK2 and Aurora B kinases in mitotic regulation. *Molecular Biology of the Cell*, 22(15):2680–2689, jun 2011. ISSN 1059-1524. doi: 10.1091/mbc.e10-11-0915.

T. Perriches and M. R. Singleton. Structure of yeast kinetochore Ndc10 DNA-binding domain reveals unexpected evolutionary relationship to tyrosine recombinases. *Journal of Biological Chemistry*, 287(7):5173–5179, 2012. ISSN 00219258. doi: 10.1074/jbc.C111.318501.

A. Petrovic, S. Pasqualato, P. Dube, V. Krenn, S. Santaguida, D. Cittaro, S. Monzani, L. Massimiliano, J. Keller, A. Tarricone, A. Maiolica, H. Stark, and A. Musacchio. The MIS12 complex is a protein interaction hub for outer kinetochore assembly. *Journal of Cell Biology*, 190(5):835–852, 2010. ISSN 00219525. doi: 10.1083/jcb.201002070.

E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13):1605–1612, oct 2004. ISSN 0192-8651. doi: 10.1002/jcc.20084.

A. L. Pidoux, E. S. Choi, J. K. R. Abbott, X. Liu, A. Kagansky, A. G. Castillo, G. L. Hamilton, W. Richardson, J. Rappsilber, X. He, and R. C. Allshire. Fission Yeast Scm3: A CENP-A Receptor Required for Integrity of Subkinetochore Chromatin. *Molecular Cell*, 33(3):299–311, 2009. ISSN 10972765. doi: 10.1016/j.molcel.2009.01.019.

L. I. Pietrasanta, D. Thrower, W. Hsieh, S. Rao, O. Stemmann, J. Lechner, J. Carbon, and H. Hansma. Probing the Saccharomyces Cerevisiae Centromeric DNA (CEN DNA)-binding Factor 3 (CBF3) Kinetochore Complex by Using Atomic Force Microscopy. *Proceedings of the National Academy of Sciences*, 96(7):3757–3762, 1999. ISSN 0027-8424. doi: 10.1073/pnas.96.7.3757.

A. F. Pluta, A. M. Mackay, A. M. Ainsztein, I. G. Goldberg, and W. C. Earnshaw. The Centromere: Hub of Chromosomal Activities. *Science*, 270(5242):1591–1594, 1995. ISSN 0036-8075. doi: 10.1126/science.270.5242.1591.

A. Poddar, J. A. Daniel, J. R. Daum, and D. J. Burke. Differential Kinetochore Requirements for Establishment and Maintenance of the Spindle Checkpoint Are Dependent on the Mechanism of Checkpoint Activation in Saccharomyces cerevisiae. *Cell Cycle*, 3(2):192–199, 2004. doi: 10.4161/cc.3.2.635.

I. Pot, V. Measday, B. Snydsman, G. Cagney, S. Fields, T. N. Davis, E. G. D. Muller, and P. Hieter. Chl4p and Iml3p Are Two New Members of the Budding Yeast Outer Kinetochore. *Molecular Biology of the Cell*, 14(2):460–476, dec 2002. ISSN 1059-1524. doi: 10.1091/mbc.e02-08-0517.

C. Prodromou, G. Siligardi, R. O'Brien, D. N. Woolfson, L. Regan, B. Panaretou, J. E. Ladbury, P. W. Piper, and L. H. Pearl. Regulation of Hsp90 ATPase activity by tetratricopeptide repeat (TPR)-domain co-chaperones. *The EMBO Journal*, 18(3): 754–762, feb 1999. ISSN 0261-4189. doi: 10.1093/emboj/18.3.754.

A. Punjani, J. L. Rubinstein, D. J. Fleet, and M. A. Brubaker. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, 14(3): 290–296, feb 2017.

A. Purvis and M. R. Singleton. Insights into kinetochore-DNA interactions from the structure of Cep3Δ. *EMBO Reports*, 9(1):56–62, 2008. ISSN 1469221X. doi: 10.1038/sj.embor.7401139.

F. Rago, K. E. Gascoigne, and I. M. Cheeseman. Distinct organization and regulation of the outer kinetochore KMN network downstream of CENP-C and CENP-T. *Current Biology*, 25(5):671–677, 2015. ISSN 09609822. doi: 10.1016/j.cub.2015.01.059.

K. Richter, S. Walter, and J. Buchner. The Co-chaperone Sba1 Connects the ATPase Reaction of Hsp90 to the Progression of the Chaperone Cycle. *Journal of Molecular Biology*, 342(5):1403–1413, 2004. ISSN 0022-2836. doi: https://doi.org/10.1016/j.jmb.2004.07.064.

I. D. Russell, A. S. Grancell, and P. K. Sorger. The Unstable F-box Protein p58-Ctf13 Forms the Structural Core of the CBF3 Kinetochore Complex. *The Journal of Cell Biology*, 145(5):933–950, 1999. ISSN 0021-9525. doi: 10.1083/jcb.145.5.933.

C. J. Russo and L. A. Passmore. Progress towards an optimal specimen support for electron cryomicroscopy. *Current Opinion in Structural Biology*, 37:81–89, 2016. ISSN 0959-440X. doi: https://doi.org/10.1016/j.sbi.2015.12.007.

P. Sahasrabudhe, J. Rohrberg, M. M. Biebl, D. A. Rutz, and J. Buchner. The Plasticity of the Hsp90 Co-chaperone System. *Molecular Cell*, 67(6):947–961.e5, sep 2017. ISSN 1097-2765. doi: 10.1016/j.molcel.2017.08.004.

L. Sanchez-Pulido, A. L. Pidoux, C. P. Ponting, and R. C. Allshire. Common Ancestry of the CENP-A Chaperones Scm3 and HJURP. *Cell*, 137(7):1173–1174, 2009. ISSN 00928674. doi: 10.1016/j.cell.2009.06.010.

S. Santaguida and A. Amon. Short- and long-term effects of chromosome mis-segregation and aneuploidy. *Nature Reviews Molecular Cell Biology*, 16(8):473–485, 2015. ISSN 14710080. doi: 10.1038/nrm4025.

S. Santaguida and A. Musacchio. The life and miracles of kinetochores. *The EMBO journal*, 28(17):2511–31, 2009. ISSN 1460-2075. doi: 10.1038/emboj.2009.173.

T. Schalch and F. A. Steiner. Structure of centromere chromatin: from nucleosome to chromosomal architecture. *Chromosoma*, 126(4):443–455, aug 2017. ISSN 1432-0886. doi: 10.1007/s00412-016-0620-7.

P. Schjerling and S. Holmberg. Comparative Amino Acid Sequence Analysis of the C6 Zinc Cluster Family of Transcriptional Regulators. *Nucleic Acids Research*, 24(23): 4599–4607, dec 1996. ISSN 0305-1048. doi: 10.1093/nar/24.23.4599.

A. Schleiffer, M. Maier, G. Litos, F. Lampert, P. Hornung, K. Mechtler, and S. Westermann. CENP-T proteins are conserved centromere receptors of the Ndc80 complex. *Nature Cell Biology*, 14(6):604–613, 2012. ISSN 14657392. doi: 10.1038/ncb2493.

F. Schmitzberger and S. C. Harrison. RWD domain: A recurring module in kinetochore architecture shown by a Ctf19-Mcm21 complex structure. *EMBO Reports*, 13(3): 216–222, 2012. ISSN 1469221X. doi: 10.1038/embor.2012.1.

B. A. Schulman, A. C. Carrano, P. D. Jeffrey, Z. Bowen, E. R. Kinnucan, M. S. Finnin, S. J. Elledge, J. W. Harper, M. Pagano, and N. P. Pavletich. Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. *Nature*, 408(6810): 381–386, 2000. ISSN 00280836. doi: 10.1038/35042620.

L. B. Sheard, X. Tan, H. Mao, J. Withers, G. Ben-Nissan, T. R. Hinds, Y. Kobayashi, F. F. Hsu, M. Sharon, J. Browse, S. Y. He, J. Rizo, G. A. Howe, and N. Zheng. Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor. *Nature*, 468(7322):400–407, 2010. ISSN 00280836. doi: 10.1038/nature09430.

M. Shivaraju, R. Camahort, M. Mattingly, and J. L. Gerton. Scm3 is a centromeric nucleosome assembly factor. *Journal of Biological Chemistry*, 286(14):12016–12023, 2011. ISSN 00219258. doi: 10.1074/jbc.M110.183640.

F. J. Sigworth. Principles of cryo-EM single-particle image processing. *Microscopy*, 65 (1):57–67, dec 2015. ISSN 2050-5698. doi: 10.1093/jmicro/dfv370.

P. K. Sorger, F. F. Severin, and A. A. Hyman. Factors required for the binding of reassembled yeast kinetochores to microtubules in vitro. *Journal of Cell Biology*, 127(4):995–1008, 1994. ISSN 00219525. doi: 10.1083/jcb.127.4.995.

D. Sousa and N. Grigorieff. Ab initio resolution measurement for single particle structures. *Journal of Structural Biology*, 157(1):201–210, 2007. ISSN 1047-8477. doi: https://doi.org/10.1016/j.jsb.2006.08.003.

F. Spencer and P. Hieter. Centromere DNA mutations induce a mitotic delay in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences*, 89(19): 8908–8912, 1992. ISSN 0027-8424. doi: 10.1073/pnas.89.19.8908.

H. Stark. Chapter Five - GraFix: Stabilization of Fragile Macromolecular Complexes for Single Particle Cryo-EM. In G. J. B. T. M. i. E. Jensen, editor, *Cryo-EM Part A Sample Preparation and Data Collection*, volume 481, pages 109–126. Academic Press, 2010. ISBN 0076-6879. doi: https://doi.org/10.1016/S0076-6879(10)81005-5.

P. Steensgaard, M. Garrè, I. Muradore, P. Transidico, E. A. Nigg, K. Kitagawa, W. C. Earnshaw, M. Faretta, and A. Musacchio. Sgt1 is required for human kinetochore assembly. *EMBO reports*, 5(6):626–631, jun 2004. ISSN 1469-221X. doi: 10.1038/sj.embor.7400154.

F. A. Steiner and S. Henikoff. Holocentromeres are dispersed point centromeres localized at transcription factor hotspots. *eLife*, 3:e02025, apr 2014. ISSN 2050-084X. doi: 10.7554/eLife.02025.

O. Stemmann and J. Lechner. The Saccharomyces cerevisiae kinetochore contains a cyclin-CDK complexing homologue, as identified by in vitro reconstitution. *The EMBO Journal*, 15(14):3611–3620, 1996. doi: 10.1002/j.1460-2075.1996.tb00730.x.

O. Stemmann, A. Neidig, T. Köcher, M. Wilm, and J. Lechner. Hsp90 enables Ctf13p/Skp1p to nucleate the budding yeast kinetochore. *Proceedings of the National Academy of Sciences*, 99(13):8585–8590, 2002. ISSN 0027-8424. doi: 10.1073/pnas.082223899.

B. M. Stern and A. W. Murray. Lack of tension at kinetochores activates the spindle checkpoint in budding yeast. *Current Biology*, 11(18):1462–1467, 2001. ISSN 0960-9822. doi: https://doi.org/10.1016/S0960-9822(01)00451-1.

S. Stoler, K. Rogers, S. Weitze, L. Morey, M. Fitzgerald-Hayes, and R. E. Baker. Scm3, an essential Saccharomyces cerevisiae centromere protein required for G2/M progression and Cse4 localization. *Proceedings of the National Academy of Sciences*, 104(25):10571–10576, 2007. ISSN 0027-8424. doi: 10.1073/pnas.0703178104.

A. V. Strunnikov, J. Kingsbury, and D. Koshland. CEP3 encodes a centromere protein of Saccharomyces cerevisiae. *Journal of Cell Biology*, 128(5):749–760, mar 1995. ISSN 0021-9525. doi: 10.1083/jcb.128.5.749.

X. Tan, L. I. A. Calderon-Villalobos, M. Sharon, C. Zheng, C. V. Robinson, M. Estelle, and N. Zheng. Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature*, 446(7136):640–645, 2007. ISSN 14764687. doi: 10.1038/nature05731.

G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, and S. J. Ludtke. EMAN2: An extensible image processing suite for electron microscopy. *Journal of Structural Biology*, 157(1):38–46, 2007. ISSN 1047-8477. doi: https://doi.org/10.1016/j.jsb.2006.05.009.

P. A. Tavormina and D. J. Burke. Cell cycle arrest in cdc20 mutants of Saccharomyces cerevisiae is independent of Ndc10p and kinetochore function but requires a subset of spindle checkpoint genes. *Genetics*, 148(4):1701–1713, 1998. ISSN 00166731.

A. N. Temiz, P. V. Benos, and C. J. Camacho. Electrostatic hot spot on DNA-binding domains mediates phosphate desolvation and the pre-organization of specificity determinant side chains. *Nucleic Acids Research*, 38(7):2134–2144, jan 2010. ISSN 0305-1048. doi: 10.1093/nar/gkp1132.

J. Thakur, P. B. Talbert, and S. Henikoff. Inner kinetochore protein interactions with regional centromeres of fission yeast. *Genetics*, 201(2):543–561, 2015. ISSN 19432631. doi: 10.1534/genetics.115.179788.

S. L. Thompson, S. F. Bakhoum, and D. A. Compton. Mechanisms of Chromosomal Instability. *Current Biology*, 20(6):R285–R295, 2010. ISSN 0960-9822. doi: https://doi.org/10.1016/j.cub.2010.01.034.

J. F. Tien, N. T. Umbreit, D. R. Gestaut, A. D. Franck, J. Cooper, L. Wordeman, T. Gonen, C. L. Asbury, and T. N. Davis. Cooperation of the Dam1 and Ndc80 kinetochore complexes enhances microtubule coupling and is regulated by aurora B. *Journal of Cell Biology*, 189(4):713–723, 2010. ISSN 00219525. doi: 10.1083/jcb.200910142.

K. Torok and D. R. Trentham. Mechanism of2-Chloro-(.epsilon.-amino-Lys75)-[6-[4-(N,N-diethylamino)phenyl]-1,3,5-triazin-4-yl]calmodulin Interactions with Smooth Muscle Myosin Light Chain Kinase and Derived Peptides. *Biochemistry*, 33(43):12807–12820, jan 1994. ISSN 0006-2960. doi: 10.1021/bi00209a012.

C. Toyoshima and N. Unwin. Contrast transfer for frozen-hydrated specimens: Determination from pairs of defocused images. *Ultramicroscopy*, 25(4):279–291, 1988. ISSN 0304-3991. doi: https://doi.org/10.1016/0304-3991(88)90003-4.

A. Traven, B. Jelicic, and M. Sopta. Yeast Gal4: a transcriptional paradigm revisited. *EMBO reports*, 7(5):496–499, may 2006. ISSN 1469-221X. doi: 10.1038/sj.embor.7400679.

J. D. Tytell and P. K. Sorger. Analysis of kinesin motor function at budding yeast kinetochores . *Journal of Cell Biology*, 172(6):861–874, mar 2006. ISSN 0021-9525. doi: 10.1083/jcb.200509101.

B. K. Vig. Sequence of centromere separation: Role of centromeric heterochromatin. *Genetics*, 102(4):795–806, 1982. ISSN 00166731.

M. Vleugel, E. Hoogendoorn, B. Snel, and G. J. Kops. Evolution and Function of the Mitotic Checkpoint. *Developmental Cell*, 23(2):239–250, 2012. ISSN 15345807. doi: 10.1016/j.devcel.2012.06.013.

R. H. Wade. A brief look at imaging and contrast transfer. *Ultramicroscopy*, 46(1): 145–156, 1992. ISSN 0304-3991. doi: https://doi.org/10.1016/0304-3991(92)90011-8.

C. E. Walczak, S. Cai, and A. Khodjakov. Mechanisms of chromosome behaviour during mitosis. *Nature Reviews Molecular Cell Biology*, 11(2):91–102, 2010. ISSN 14710072. doi: 10.1038/nrm2832.

J. Wang, N. Suzuki, Y. Nishida, and T. Kataoka. Analysis of the function of the 70-kilodalton cyclase-associated protein (CAP) by using mutants of yeast adenylyl cyclase defective in CAP binding. *Molecular and Cellular Biology*, 13(7):4087–4097, jul 1993. doi: 10.1128/MCB.13.7.4087.

M. N. Waxham, A. L. Tsai, and J. A. Putkey. A mechanism for calmodulin (CaM) trapping by CaM-kinase II defined by a family of CaM-binding peptides. *Journal of Biological Chemistry*, 273(28):17579–17584, 1998. ISSN 00219258. doi: 10.1074/jbc.273.28.17579.

B. A. Weaver and D. W. Cleveland. Does aneuploidy cause cancer? *Current Opinion in Cell Biology*, 18(6):658–667, 2006. ISSN 09550674. doi: 10.1016/j.ceb.2006.10.002.

R. R. Wei, P. K. Sorger, and S. C. Harrison. Molecular organization of the Ndc80 complex, an essential kinetochore component. *Proceedings of the National Academy of Sciences*, 102(15):5363–5367, 2005. ISSN 0027-8424. doi: 10.1073/pnas.0501168102.

R. R. Wei, J. R. Schnell, N. A. Larsen, P. K. Sorger, J. J. Chou, and S. C. Harrison. Structure of a Central Component of the Yeast Kinetochore: The Spc24p/Spc25p Globular Domain. *Structure*, 14(6):1003–1009, 2006. ISSN 09692126. doi: 10.1016/j.str.2006.04.007.

J. R. Weir, A. C. Faesen, K. Klare, A. Petrovic, F. Basilico, J. Fischböck, S. Pentakota, J. Keller, M. E. Pesenti, D. Pan, D. Vogt, S. Wohlgemuth, F. Herzog, and A. Musacchio. Insights from biochemical reconstitution into the architecture of human kinetochores. *Nature*, 537:249–253, aug 2016.

S. Westermann, H.-W. Wang, A. Avila-Sakar, D. G. Drubin, E. Nogales, and G. Barnes. The Dam1 kinetochore ring complex moves processively on depolymerizing microtubule ends. *Nature*, 440(7083):565–569, 2006. ISSN 1476-4687. doi: 10.1038/nature04409.

R. Wevrick and H. F. Willard. Long-range organization of tandem arrays of alpha satellite DNA at the centromeres of human chromosomes: high-frequency array-length polymorphism and meiotic stability. *Proceedings of the National Academy of Sciences*, 86(23):9394–9398, 1989. ISSN 0027-8424. doi: 10.1073/pnas.86.23.9394.

G. Wieland, S. Orthaus, S. Ohndorf, S. Diekmann, and P. Hemmerich. Functional Complementation of Human Centromere Protein A (CENP-A) by Cse4p from Saccharomyces cerevisiae. *Molecular and Cellular Biology*, 24(15):6620–6630, 2004. ISSN 0270-7306. doi: 10.1128/MCB.24.15.6620-6630.2004.

P. A. Wigge and J. V. Kilmartin. The Ndc80p complex from Saccharomyces cerevisiae contains conserved centromere components and has a function in chromosome segregation. *Journal of Cell Biology*, 152(2):349–360, 2001. ISSN 00219525. doi: 10.1083/jcb.152.2.349.

O. Willhoft, R. Kerr, D. Patel, W. Zhang, C. Al-Jassar, T. Daviter, S. H. Millson, K. Thalassinos, and C. K. Vaughan. The crystal structure of the Sgt1-Skp1 complex: The link between Hsp90 and both SCF E3 ubiquitin ligases and kinetochores. *Scientific Reports*, 7:41626, 2017. ISSN 20452322. doi: 10.1038/srep41626.

J. S. Williams, T. Hayashi, M. Yanagida, and P. Russell. Fission Yeast Scm3 Mediates Stable Assembly of Cnp1/CENP-A into Centromeric Chromatin. *Molecular Cell*, 33 (3):287–298, 2009. ISSN 10972765. doi: 10.1016/j.molcel.2009.01.017.

H. Xiao, F. Wang, J. Wisniewski, A. K. Shaytan, R. Ghirlando, P. C. Fitzgerald, Y. Huang, D. Wei, S. Li, D. Landsman, A. R. Panchenko, and C. Wu. Molecular basis of CENP-C association with the CENP-A nucleosome at yeast centromeres. *Genes and Development*, 31(19):1958–1972, 2017. ISSN 15495477. doi: 10.1101/gad.304782.117.

Y. Yamagishi, T. Sakuno, Y. Goto, and Y. Watanabe. Kinetochore composition and its function: Lessons from yeasts. *FEMS Microbiology Reviews*, 38(2):185–200, 2014. ISSN 01686445. doi: 10.1111/1574-6976.12049.

K. Yan, Z. Zhang, J. Yang, S. H. McLaughlin, and D. Barford. Architecture of the CBF3–centromere complex of the budding yeast kinetochore. *Nature Structural and Molecular Biology*, 25:1103–1110, 2018. ISSN 15459985. doi: 10.1038/s41594-018-0154-1.

J. W. Yang, C. Pendon, J. Yang, N. Haywood, A. Chand, and W. R. A. Brown. Human mini-chromosomes with minimal centromeres. *Human Molecular Genetics*, 9(12): 1891–1902, jul 2000. ISSN 0964-6906.

Y. Yang, F. Wu, T. Ward, F. Yan, Q. Wu, Z. Wang, T. McGlothen, W. Peng, T. You, M. Sun, T. Cui, R. Hu, Z. Dou, J. Zhu, W. Xie, Z. Rao, X. Ding, and X. Yao. Phosphorylation of HsMis13 by Aurora B kinase is essential for assembly of functional kinetochore. *Journal of Biological Chemistry*, 283(39):26726–26736, 2008. ISSN 00219258. doi: 10.1074/jbc.M804207200.

Y. B. Yurov, S. G. Vorsanova, T. Liehr, A. D. Kolotii, and I. Y. Iourov. X chromosome aneuploidy in the Alzheimer's disease brain. *Molecular cytogenetics*, 7(1):20, 2014. ISSN 1755-8166.

K. Zhang. Gctf: Real-time CTF determination and correction. *Journal of Structural Biology*, 193(1):1–12, 2016. ISSN 1047-8477. doi: https://doi.org/10.1016/j.jsb.2015.11.003.

M. Zhang, M. Botër, K. Li, Y. Kadota, B. Panaretou, C. Prodromou, K. Shirasu, and L. H. Pearl. Structural and functional coupling of Hsp90- and Sgt1-centred multi-protein complexes. *The EMBO Journal*, 27(20):2789–2798, oct 2008. ISSN 0261-4189. doi: 10.1038/emboj.2008.190.

W. Zhang, N. Lukoyanova, S. Miah, J. Lucas, and C. K. Vaughan. Insights into Centromere DNA Bending Revealed by the Cryo-EM Structure of the Core Centromere Binding Factor 3 with Ndc10. *Cell Reports*, 24(3):744–754, jul 2018. ISSN 2211-1247. doi: 10.1016/j.celrep.2018.06.068.

R. Zhao, M. Davey, Y.-C. Hsu, P. Kaplanek, A. Tong, A. B. Parsons, N. Krogan, G. Cagney, D. Mai, J. Greenblatt, C. Boone, A. Emili, and W. A. Houry. Navigating the Chaperone Network: An Integrative Map of Physical and Genetic Interactions Mediated by the Hsp90 Chaperone. *Cell*, 120(5):715–727, jan 2005. ISSN 0092-8674. doi: 10.1016/j.cell.2004.12.024.

S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, and D. A. Agard. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods*, 14(4):331–332, feb 2017.

J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. H. Hagen, E. Lindahl, and S. H. W. Scheres. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife*, 7:e42166, 2018. ISSN 2050-084X. doi: 10.7554/eLife.42166.

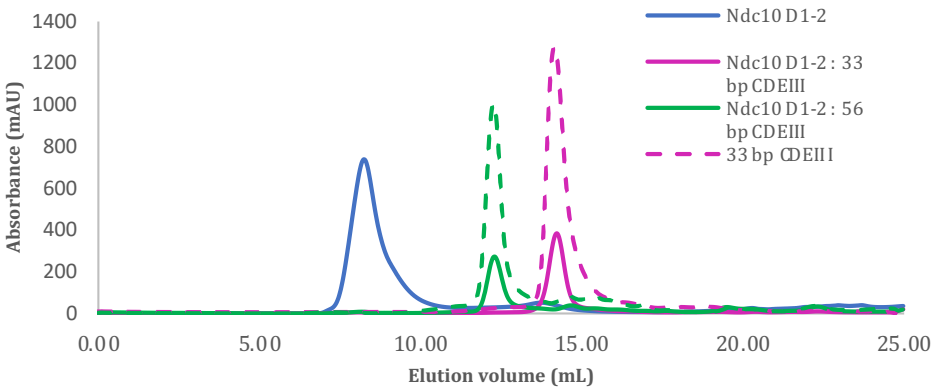
Appendix A

Appendix

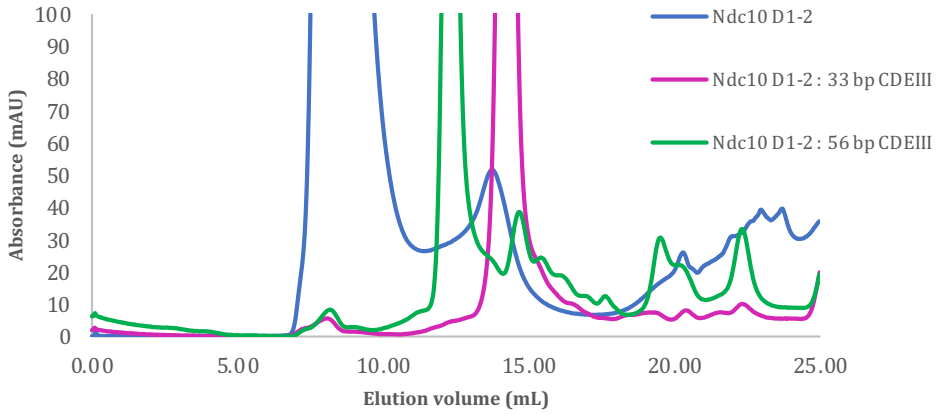
CEN1	GTCACATGACATAATAATAAATAATTTTAAAAATATAAAATATTTTTAATAGTTTTTAAA TATTTTACAGTTTATTTTTTAAATTTATTTATATGTTTTTGTTTTCCGAAGCAGTCAA
CEN2	ATCATGTGACTTATTTATTTAATTATTATTAAGTAAAAAAGATTTTCTATTTAAATTTATT AATTAATTTTTTTTCTTAAATAATTATTTTATGTTTTTGTTTTCCGAAAAAGAAAA
CEN3	GTCACATGATGATATTTGATTTTATTATATTTTTAAAAAAAGTAAAAAATAAAAAGTAG TTTATTTTTAAAAAATAAAATTTAAAATATTAGTGTATTTGATTTCCGAAAGTTAAAA
CEN4	GTCACATGCTTATAATCAACTTTTTTAAAAATTTAAAATACTTTTTTATTTTTTATTTTT AAACATAAATGAAATAATTTATTTATTGTTTATGATTACCGAAACATAAAA
CEN5	ATCACGTGCTTTTTAAAAAATATAAATTTAATTTCATTTTCTATTTCAATATTTATTAAAT AAAAAATTTGAAAAATATATAAAAATTGTAGCAGTATTAGATTTCCGAAAAGAAAAA
CEN6	ATCACGTGCTATAAAAATAATTATAATTTAAATTTTTTAATATAAATATATAAATTAAAA ATAGAAAGTAAAAAAAGAAATTAAAGAAAAAATAGTTTTTGTTTTCCGAAGATGTAAA
CEN7	ATCACGTGTTATATTTACTATATAAAAATTCAATAAATAAAAAGTTAGAAGATAAAAAT TATATTATACATATTTTTATTTTTATTATAATTTTTGTTTTTGCCTTCCGAAAAGAAAAT
CEN8	ATCACATGACTAATAATTCTTTTAATTTTAATTAATTTAATAAAATTAAAATAATATATAT ACTAAATTGTTTATTAAAAATGATTAAACATTGGGTTTTGTGTTCCGAACTTAGAAA
CEN9	TTCACGTGAAAATTTTTATATTTTTAATTAAATTTTTATAATATTATAAATTATTATAATA TTGATATTTAAAATTAAAAACAAATTATTAATGGTTTTGTTTTCCGAAATGTTTTT
CEN10	ATCACGTGTTAAATAATTAATTTACTTTAAAATTTATTTTTTAATATAAAATATTTATTCT TTTTATTTAAAAATAAAAAACACAAAAAAACAATGTTTATGATTTCCGAACCTAAATA
CEN11	GTCACATGATAAAAACATATTTAAAATTTTAAAAAAATTAATTTTCAAAATAAATTTATT ATATTTTTTTAATTACATAATCATAAAAATAAATGTTCATGATTTCCGAACGTATAAA
CEN12	ATCACGTGTAATAAATATTATTAAAAAGTTTATTAAAATAAAATAATAATTTAAATTACT ATTTTTAAATAAGTTTTATTTTTTAATAACACTATTGTATTTGTTATCCGAACAATAAAA
CEN13	ATCACATGACTACCTAACAAAATATTTATTTTTCTTTTTTAATATTTGAAAATACTAAAAT ATTTTTGTTGTTTTTTGAAAAAAGGATTTTTAATGTGTATGCGTTCCGAACTTTAAAT
CEN14	GTCACGTGCAGCTTTTTAAAAATATTTTAAAACATTTTAAAAAATATACATTTTTTTATTA TTTTTTTATATATTAATGTTAAAATTTATTTATGTATTTGTCTTCCGAAAAGTAAAA
CEN15	ATCACGTGAACTTATTTTGCATTTAAAAAAAAGTAAAAACTATTTGCTAAAATATATTTT TTTAATTTTTAAAAATAATGTTTTAATTATTTAATGTATATGACTTCCGAAAAATATAT
CEN16	ATCACATGATATATTTTTTATTTTTAATTTTTTTTAATTATAAAAATAATTTTTTTCTTT AAATTAAACAAAAATAAAAAATTGTTTTTTGTTGGTTAAGATTTCCGAAAATAGAAA

Fig. A.1 **CEN sequences.** List of all 16 CEN DNA sequences from *S. cerevisiae* obtained from uniprot.

A



B



C

Sample	Peak elution volume (mL)	Elution volume range (mL)
Ndc10 D1-2	13.73	11.15 – 15.35
Ndc10 D1-2 - 33 bp	13.95	13.75 – 15.95
Ndc10 D1-2 - 56 bp	13.95	12.75 – 14.75
33 bp CDEIII	14.15	13.51 – 16.00
56 bp CDEIII	12.27	11.70 - 13.70

Fig. A.2 **Gel filtration analysis of Ndc10 binding in S200 buffer.** (A) Comparison of gel filtration chromatograms of Ndc10 domains 1 and 2 alone, with 33 bp and 56 bp CDEIII sequences. The protein and DNA incubation was conducted in gel filtration buffer. (B) Re-scaled view of (A) for easier peak comparison. (C) List of the peak positions for each sample.